

Electronic Thesis and Dissertation Repository

---

5-25-2011 12:00 AM

## Acquisition de la liaison et de l'enchaînement en français L2 : le rôle de la fréquence

Nadine Odette de Moras, *The University of Western Ontario*

Supervisor: Jeff Tennant, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in French

© Nadine Odette de Moras 2011

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [French Linguistics Commons](#)

---

### Recommended Citation

de Moras, Nadine Odette, "Acquisition de la liaison et de l'enchaînement en français L2 : le rôle de la fréquence" (2011). *Electronic Thesis and Dissertation Repository*. 176.  
<https://ir.lib.uwo.ca/etd/176>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

Acquisition de la liaison et de l'enchaînement en français L2 :

Le rôle de la fréquence

(Spine title: Acquisition de la liaison et de l'enchaînement en français L2)  
(Thesis format: monograph)

by

Nadine de Moras

Graduate Program in French Studies

A thesis submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies  
The University of Western Ontario  
London, Ontario, Canada

© Nadine de Moras 2011

**CERTIFICATE OF EXAMINATION**

Supervisor

Examiners

\_\_\_\_\_  
Dr. Jeff Tennant

\_\_\_\_\_  
Dr. Joyce Bruhn de Garavito

Supervisory Committee

\_\_\_\_\_  
Dr. Julie Byrd Clark

\_\_\_\_\_  
Dr. François Poiré

\_\_\_\_\_  
Dr. Alain Thomas

La thèse de

**Nadine Odette de Moras**

s'intitule:

Acquisition de la liaison et de l'enchaînement en français L2: le rôle de la fréquence

is accepted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

\_\_\_\_\_  
Date

\_\_\_\_\_  
Chair of the Thesis Examination Board

## Abstract

Après un survol des recherches générales en acquisition des enfants et adultes L1 et L2 dans le cadre du Modèle Basé sur l'Usage, en acquisition des liaisons des enfants et des adultes francophones, et des recherches sur les productions de liaisons des étudiants L2 de niveau intermédiaire, nous avons analysé le processus d'acquisition de la liaison.

Les étudiants anglophones de Mastro Monaco (2000), Thomas (2002) et Howard (2005) ont produit plus de 91% de liaisons obligatoires, pourcentages qui suggèrent que les anglophones maîtrisent les liaisons françaises. Une analyse approfondie des méthodes de recherches et de codage explique ces taux élevés.

Nous avons créé un texte que 20 francophones et 37 anglophones ont lu. Les francophones majoritaires ont produit 95.6% de liaisons obligatoires, les francophones minoritaires en ont produit 85.5% et les étudiants anglophones 60.7%.

Nous avons testé les effets des répétitions, des corrections et des explications dans l'enseignement de la phonétique et une semaine plus tard les effets combinés de tous ces éléments. Au post-test les étudiants ont produit 69.5% de liaisons obligatoires, ce qui représente des progrès visibles. Le groupe qui a fait le plus de progrès est le groupe ayant eu les répétitions. Le groupe qui en a fait le moins est celui qui a reçu le plus d'explications théoriques.

Les francophones utilisent des informations lexicales et fréquentielles : la fréquence du mot1 et du mot2, la fréquence de cooccurrence, la fréquence et la productivité des structures syntaxiques ainsi que des informations phonétiques et prosodiques, dans leur analyse multidimensionnelle de la liaison. Les apprenants L2 utilisent ces mêmes informations, mais ne semblent pas avoir assez d'éléments dans leur banque de données pour pouvoir traiter les informations avec autant de précision et d'efficacité que les francophones.

Nous avons aussi étudié les effets d'autres éléments sociolinguistiques pouvant rentrer en compte dans la production de liaisons et d'enchaînements des apprenants : la L1, le type d'enseignement reçu à l'école, le temps passé en pays/région francophone, le nombre d'activités en français, et le nombre d'années passées à étudier le français.

## Mots-clés (en français)

Liaisons; enchaînements; Phonétique française; Prononciation; Modèle Basé sur l'Usage; Fréquence; Acquisition des langues secondes; Enseignement des langues secondes ; Langue orale; Enseignement de la prononciation; Analyse comparative;

## Mots-clés (en anglais)

Consonant linking; Liaisons; French Phonetics; Pronunciation; Usage-Based Model; Frequency; Second Language Learning; Second Language Teaching; Oral Language; teaching pronunciation; Comparative Analysis;

## Remerciements

Je voudrais remercier toutes les personnes qui ont contribué à cette thèse. Je voudrais tout d'abord remercier mon directeur de thèse, Jeff Tennant, qui m'a guidée dans ma réflexion, mes recherches et toutes les étapes du codage et de l'analyse des données.

Je remercie les membres du comité et leur aide précieuse, en particulier Alain Thomas, qui a largement contribué améliorer cette thèse, dans le fond, comme dans la forme.

Je voudrais aussi remercier Suzanne Majhanovich qui a été une source d'inspiration, d'information et de réflexion depuis mes premiers cours de maîtrise.

Je remercie aussi tous les professeurs que j'ai eus en maîtrise et en doctorat, mais aussi ceux que j'ai eu avant, ainsi que mes collègues de travail qui m'ont encouragée à continuer mes études, tout en travaillant à plein temps.

Je remercie mon mari qui m'a soutenue et m'a aidée à terminer cette thèse. Enfin, je remercie mes enfants pour leur patience et leur compréhension qui eux aussi ont été une source d'inspiration et de motivation.

## TABLE DES MATIÈRES

CERTIFICATE OF EXAMINATION	ii
Abstract	iii
Mots-clés	iv
Remerciements	v
Table des matières	vi
Liste des tableaux	ix
Liste des figures et graphiques	xii
Liste des Annexes	xiii
Préface	xiv
<b>Introduction</b>	<b>1</b>
<b>Chapitre 1 : les recherches sur les liaisons et le Modèle Basé sur l'Usage</b>	<b>17</b>
Introduction	18
1. Recherches sur l'enseignement des langues, le rôle de la fréquence et le MBU	21
1.1. Survol des méthodes d'enseignement des langues	21
1.2. Recherches menant au modèle basé sur l'usage (MBU)	27
1.3. Autres études liées au Modèle Basé sur l'Usage	42
1.3.1. Le rôle de l'input : un input riche	42
1.3.2. Fréquence et évolution de la langue	47
1.3.3. Fréquence, connexionnisme et variationnisme	51
1.3.4. Recherches en reconnaissance de la parole	54
1.3.5. Le modèle basé sur l'usage et l'acquisition d'une L2	56
2. Origines, catégories et explications des liaisons	69
2.1. Les liaisons et l'évolution de la langue	75
2.2. Les catégories de liaisons	82
2.3. Les explications « classiques »	91
Une explication historique	92
Une explication syntaxique	93
Une explication morphologique	94
Une explication phonologique et prosodique	96
Une explication orthographique	99
Une explication stylistique	100
Une explication lexicale	101
2.4. Le cumul des facteurs	102
2.5. Les liaisons chez les adultes francophones	105
3. Les liaisons et le MBU	113
3.1. La fréquence des structures syntaxiques	114
3.2. La fréquence des mots monosyllabiques	118
3.3. La fréquence des phonèmes de liaisons	120
3.4. La fréquence lexicale	121
4. Les liaisons « hors-normes » et le MBU	132
4.1. Liaisons « hors-normes » et définitions	134
4.2. Cas particuliers: liaisons non-enchaînées et liaisons devant consonnes	140
4.3. Les erreurs de liaisons : explications	145
Discussion	160

<b>Chapitre 2 : L'acquisition des liaisons en L1 et en L2</b>	164
Introduction	165
1. L'acquisition des liaisons chez les enfants francophones	167
1.1. Principes généraux d'acquisition d'après le MBU	167
1.2. Lacunes dans le développement langagier de l'enfant	171
1.3. Recherches en acquisition chez les enfants L1	173
2. Les productions de liaisons des apprenants du français L2	205
2.1. Analyse de ces études et interprétation des données	217
2.1.1. Les tests utilisés	220
2.1.2. Classement des liaisons, catégorisations et définitions	226
2.1.3. Les éléments testés	228
Discussion sur les productions d'apprenants du français L2	230
<b>Chapitre 3 : Méthodologie et hypothèses</b>	235
1. La problématique	236
2. Méthodologie	237
2.1. La parole spontanée	238
2.2. La parole semi-guidée	245
2.3. La répétition	248
2.4. La lecture d'un texte à haute voix	248
2.5. Comparaison de groupes différents	251
2.6. Comparaison de productions de francophones et comparaisons de tests	253
2.7. Les comparaisons de lectures de plusieurs études	259
3. La procédure	261
4. Les sujets	262
4.1. Les sujets francophones	262
4.2. Les sujets anglophones	270
5. Les méthodes d'enseignement utilisées dans l'étude	274
6. Le texte à lire	282
7. Fréquence lexicale et banque de données	288
7.1.1. La banque de données de Lexique 3	295
7.1.2. Analyse des tableaux de Lexique 3	297
7.2.1. La banque de données de Vocabprofil	301
Texte de l'étude « décortiqué »	302
7.2.2. La fréquence des mots - discussion des données	304
7.2.3. La fréquence des mots <sup>1</sup> et des mots <sup>2</sup> de notre texte et le codage	308
8. La fréquence et la productivité des structures syntaxiques	309
9. La fréquence des phonèmes de liaison	316
9.1. Le nombre de structures de chacune des consonnes de liaison	316
9.2. Consonne de liaison et structure morphosyntaxique et sémantique	318
9.3. La fréquence des consonnes de liaisons	321
9.4. Les contextes phonétiques de la consonne de liaison	325
9.4.1. Les phonèmes se trouvant juste avant la consonne de liaison	325
9.4.2. Le nombre de liaisons et de syllabes avant et après la consonne de liaison	326
9.5. La complexité des structures grapho-phonémiques	329
10. Hypothèses	334
Conclusion	342



<b>Chapitre 4 : Résultats et analyses (pourcentages)</b>	344
1. Les productions de liaisons et enchaînements des francophones	345
1.1. Les études sur les productions de francophones	345
1.2. Résultats de notre étude	347
1.2.1. Résultats des productions de francophones	348
1.2.1.1. Le sexe des sujets	348
1.2.1.2. L'âge des sujets	349
1.2.1.3. Le niveau d'étude	349
1.2.2. Résultats généraux	350
1.2.3. Les productions de liaisons obligatoires – tous francophones	356
1.2.4. Les productions de francophones et catégories grammaticales	358
1.2.5. Les productions de francophones et la fréquence lexicale	363
1.2.5.1. Analyse du mot1 avec Vocabprofil et Lexique	364
1.2.5.2. Analyse du mot2 avec Vocabprofil et Lexique	367
1.2.5.3. Analyse de cooccurrence mots1- mots2 Lexique	371
1.2.6. Les productions de francophones et les consonnes de liaison	373
Discussion sur les résultats des francophones	375
2. Les productions de liaisons et enchaînements des anglophones	384
2.1. Les résultats de notre étude	386
2.1.1. Résultats généraux	387
2.1.2. Les liaisons selon les groupes linguistiques	390
2.1.2.1. Les liaisons non-enchaînées	390
2.1.2.2. Résultats généraux par groupes linguistiques	398
2.1.2.3. Résultats par consonnes de liaison	405
2.1.3. Résultats par catégories grammaticales	408
2.1.4. Résultats par catégories fréquentielles	413
2.1.5. Résultats par groupe expérimental	418
2.1.6. Résultats par type de programme d'étude	425
2.1.7. Résultats par nombre d'années passées à étudier le français	431
2.1.8. Résultats par séjour en milieu francophone	433
2.1.9. Résultats par activités en français	435
2.1.10. Résultats selon auto-évaluation	437
Discussion	439
<b>Chapitre 5 : analyse Goldvarb et discussion</b>	445
Introduction	446
1. Analyse GoldVarb	447
2. Discussion sur les résultats obtenus	459
<b>Conclusion</b>	480
<b>Annexes</b>	498
<b>Bibliographie</b>	530
Bref CV de N. de Moras	551

## Liste des tableaux

Tableau 1.1. % de liaisons réalisées, structures syntaxiques, BDM, AD et G. (2003:1331)	88
Tableau 1.2. % de liaisons réalisées des déterminants BDM, AD et G. (2003: 1332)	89
Tableau 1.3. Liaisons avec adjectifs avec nasales, fréquence et réalisation	103
Tableau 1.4. Séquences étudiées par Sampson (2001) et fréquence de coocc. Lexique	104
Tableau 1.5. Taux de liaison dans deux corpus, Fougeron et Delais-Roussarie (2004 :2)	107
Tableau 1.6. % de réalisation de la liaison, lecture et conversation, PFC	110
Tableau 1.7. % de réalisation de la liaison, lecture et conversation par régions, PFC	111
Tableau 1.8. % de réalisation de la liaison hommes/femmes France, PFC	113
Tableau 1.9. Fréquence des structures BDM, AD et G. (2003: 1331)	116
Tableau 1.10. Classement des structures syntaxiques d'après Gadet (1989 : 52-53)	117
Tableau 1.11. % de liaisons Gadet (1989), BDM, AD et G. (2003), Mallet (2008)	118
Tableau 1.12. % de réalisation de la liaison par nombre de syllabes (d'après projet PFC)	119
Tableau 1.13. Contextes de liaisons possibles, % liaisons, Mallet (2008 : 221, 256)	120
Tableau 1.14. Propositions de distribution des CL /n/, /t/, /z/ Mallet (2008 : 205)	121
Tableau 1.15. % de réalisation de liaisons facultatives d'après différentes études	125
Tableau 1.16. Fréquence et productivité de la séquence un grand émoi, (D & L, 2008: 44)	128
Tableau 1.17. Fréquence et productivité de la séquence un grand honneur (D & L, 2008: 44)	128
Tableau 1.18. Exemple 1 de caractéristiques sociolinguistiques d'un locuteur X	130
Tableau 1.19. Exemple 2 de caractéristiques sociolinguistiques d'un locuteur Y	130
Tableau 1.20. Exemple, cumul fréquence et productivité grand honneur	131
Tableau 1.21. Exemple – cumul fréquence et productivité grand émoi	131
Tableau 1.22. Définitions des diverses erreurs avec les liaisons	139
Tableau 1.23. % de réalisation de la liaison en lecture d'après les données du projet PFC	143
Tableau 1.24. % de réalisation de la liaison en conversation données du projet PFC	144
Tableau 1.25. Séquences en haut de la hiérarchie, erreurs, Morin et Kaye (1982 : 323-24)	158
Tableau 1.26. Séquences en haut de la hiérarchie, erreurs Gadet (1989 : 56 )	159
Tableau 2.1. % d'erreurs de liaisons chez les enfants L1. Chevrot, Dugua et Fayol (2005:41)	184
Tableau 2.2. % de LO et LF, enfants L1, Chevrot, Dugua et Fayol (2005 : 43)	184
Tableau 2.3. LO enfants 2-6 ans / classe sociale, Chevrot, Dugua et Fayol (2005 : 7)	185
Tableau 2.4. Pourcentages de LF réalisées justes selon les milieux, Chabanal (2007 : 6)	187
Tableau 2.5. Liaison avec le verbe « être ». Chabanal et Embarki (2002 : 178)	188
Tableau 2.6. Résultats de l'étude transversale Dugua et Chabanal (2006 : 2)	191
Tableau 2.7. Résultats de l'étude longitudinale Dugua et Chabanal (2006 : 3)	192
Tableau 2.8. 5 études de cas : L correctes et erreurs Chev., Chab. et Dug. (2007 : 112)	195
Tableau 2.9. LO et LF, et erreurs enfants 40/50 mois, Chev., Chab. et Dug. (2007 : 117)	197
Tableau 2.10. Erreurs de liaisons, enfants 4-12 ans, Pallaud et Savelli (2001 : 124-127)	201
Tableau 2.11. Erreurs, séquences en haut de la hiérarchie, Pallaud et Savelli (2001 :124-7)	201
Tableau 2.12. % LO d'après les contextes syntaxiques Mastro Monaco (2000 : 146)	206
Tableau 2.13. % LO selon le contexte syntaxique. Howard (2005 : 12)	212
Tableau 2.14. L anglophones, Mastro Monaco (2000), Thomas (2002) et Howard (2005)	216
Tableau 3.1. Les sessions de l'étude.	261
Tableau 3.2. Renseignements du questionnaire des sujets francophones (majoritaires)	268
Tableau 3.3. Renseignements du questionnaire des sujets francophones (minoritaires)	268
Tableau 3.4. Horaire des activités de l'étude expérimentale	279
Tableau 3.5. Fréquence mot1 Lexique 3	295
Tableau 3.6. Fréquence mot2 Lexique 3	295

Tableau 3.7. Fréquence cooccurrence Lexique 3	296
Tableau 3. 8. Fréquence des mots avec Vocabprofil	302
Tableau 3. 9. Séquences « difficiles à traiter » d’après la fréquence avec Vocabprofil	305
Tableau 3.10. Contextes liants et liaisons réalisées selon les phonèmes. Mallet (2008:218)	318
Tableau 3.11. Fréquence des articles dans les films et livres avec Lexique	322
Tableau 3.12. Distribution des consonnes de liaison réalisées (LR) Mallet (2008 : 256)	322
Tableau 3.13. Fréquence des liaisons réalisées et non réalisées. Mallet (2008 : 213)	323
Tableau 3.14. Fréquence des phonèmes de liaisons d’après les catégories syntaxiques	323
Tableau 3.15. Rôle du contexte phonétique précédant la liaison	326
Tableau 4.1. Productions de liaisons obligatoires par sexe	349
Tableau 4.2. Age et production de liaisons obligatoires	349
Tableau 4.3. Années d’études universitaires et production de liaisons	350
Tableau 4.4. Production de liaisons et enchaînements francophones majoritaires	354
Tableau 4.5. Production de liaisons et enchaînements francophones minoritaires	355
Tableau 4.6. Production de liaisons obligatoires tous francophones	357
Tableau 4.7. Résultats projet PFC, en ligne (notre tableau à partir des résultats en ligne)	357
Tableau 4.8. Liaisons obligatoires et catégories grammaticales tous francophones	359
Tableau 4.9. LO et catégories grammaticales, francophones majoritaires	360
Tableau 4.10. LO et catégories grammaticales, francophones minoritaires	361
Tableau 4.11. Expressions figées - Tous francophones	362
Tableau 4.12. Expressions figées - Francophones majoritaires et minoritaires	362
Tableau 4.13. Expressions figées et adjectifs, liaisons obligatoires - tous francophones	363
Tableau 4.14. Fréquence, production LO, mot1 Vocabprofil, francophones majoritaires	365
Tableau 4.15. Fréquence, production LO, mot1 Vocabprofil, francophones minoritaires	365
Tableau 4.16. Fréquence Mot1 Lexique LO – Tous francophones	366
Tableau 4.17. Fréquence Mot1 Lexique LO – Francophones majoritaires	366
Tableau 4.18. Fréquence Mot1 Lexique LO – Francophones minoritaires	366
Tableau 4.19. Fréquence, production LO, mot2 Vocabprofil - francophones majoritaires	369
Tableau 4.20. Fréquence, production LO, mot2 Vocabprofil - francophones minoritaires	369
Tableau 4.21. Fréquence Mot2 Lexique LO – Tous francophones	370
Tableau 4.22. Fréquence Mot2 Lexique LO – Francophones majoritaires	371
Tableau 4.23. Fréquence Mot2 Lexique LO – Francophones minoritaires	371
Tableau 4.24. Fréquence de cooccurrence Lexique LO – Tous francophones	372
Tableau 4.25. Fréquence de cooccurrence Lexique LO – Francophones majoritaires	372
Tableau 4.26. Fréquence de cooccurrence Lexique LO – Francophones minoritaires	373
Tableau 4.27. Productions francophones L1, PFC	374
Tableau 4.28. Consonnes de liaison et productions de liaisons obligatoires (notre étude)	375
Tableau 4.29. Production de LO, LF, LI et enchaînements francophones et anglophones	389
Tableau 4.30. Liaisons obligatoires non-enchaînées, francophones et anglophones	395
Tableau 4.31. Liaisons facultatives non-enchaînées, francophones et anglophones	395
Tableau 4.32. Liaisons non-enchaînées par occurrences et groupes	396
Tableau 4.33. LO, LF, LI non-enchaînées par groupe linguistique	401
Tableau 4.34. Liaisons obligatoires non-enchaînées par groupe linguistique	401
Tableau 4.35. Liaisons facultatives non-enchaînées par groupe linguistique	401
Tableau 4.36. Liaisons interdites non-enchaînées par groupe linguistique	402
Tableau 4.37. Productions de liaisons et enchaînements par langue maternelle	402
Tableau 4.38. Productions de liaisons obligatoires par langue maternelle	402
Tableau 4.39. Productions de liaisons facultatives par langue maternelle	402
Tableau 4.40. Productions de liaisons interdites par langue maternelle	402

Tableau 4.41. Productions d'enchaînements par langue maternelle	403
Tableau 4.42. Productions de liaisons obligatoires – consonnes de liaison – anglophones	406
Tableau 4.43. Productions de liaisons obligatoires – consonnes de liaison – bilingues	406
Tableau 4.44. Productions de liaisons obligatoires – CL – allemand et arabe	407
Tableau 4.45. Productions de LO – consonnes de liaison – cantonais	407
Tableau 4.46. Productions de LO – consonnes de liaison – tous francophones	407
Tableau 4.47. LO et consonne de liaison, tous francophones et anglophones	408
Tableau 4.48. LO par catégories grammaticales, francophones maj., min. et anglophones	412
Tableau 4.49. LO par séquences. Francophones majoritaires et anglophones	412
Tableau 4.50. Expressions figées (LO). Tous francophones et anglophones	413
Tableau 4.51. LO et fréquence du mot1 avec Vocabprofil. Tous franc. et ang.	415
Tableau 4.52. LO et fréquence du mot1 avec Lexique. Tous franc. et ang.	415
Tableau 4.53. LO et fréquence du mot2 avec Vocabprofil. Tous franc. et ang.	416
Tableau 4.54. LO et fréquence du mot2 avec Lexique. Tous franc. et ang.	416
Tableau 4.55. LO et fréquence de coocc. des mot1-mot2 avec Lexique. franc. et ang.	417
Tableau 4.56. Production de liaisons interdites	421
Tableau 4.57. Évolution du traitement de la liaison après leçon. Groupe 1	423
Tableau 4.58. Évolution du traitement de la liaison après leçon. Groupe 2	423
Tableau 4.59. Évolution du traitement de la liaison après leçon. Groupe 3	423
Tableau 4.60. Évolution du traitement de la liaison après leçon. Groupe 4	423
Tableau 4.61. Résultats par type d'école fréquentée - Liaisons obligatoires	429
Tableau 4.62. Progrès tous types de sandhi	430
Tableau 4.63. Résultats par type d'école fréquentée – enchaînements	430
Tableau 4.64. Résultats par type d'école fréquentée - liaisons facultatives	431
Tableau 4.65. Résultats par type d'école fréquentée - liaisons interdites	431
Tableau 4.66. Résultats par type d'école fréquentée - Tous types de sandhi	431
Tableau 4.67. Nombre d'années étudiées -Tous types de sandhi	432
Tableau 4.68. Liaisons obligatoires et temps passé en pays/région francophone	435
Tableau 4.69. Liaisons obligatoires et temps passé en pays/région francophone	435
Tableau 4.70. Tous types de sandhi et nombre d'activités en français	437
Tableau 4.71. Niveau estimé en prononciation (auto-évaluation) et productions de LO	438
Tableau 5.1. Réalisations LO : Francophones maj., franc. min., apprenants	446
Tableau 5.2. Réalisations LF : Francophones maj., franc. min., apprenants	447
Tableau 5.3. Réalisations LI : Francophones maj., franc. min., apprenants	448
Tableau 5.4. Réalisations enchaînements Francophones maj., franc. min., apprenants	449
Tableau 5.5. Liaisons obligatoires, francophones. Fréquence Lexique mot1	449
Tableau 5.6. Liaisons obligatoires, apprenants. Fréquence Lexique mot1	450
Tableau 5.7. Liaisons obligatoires, francophones. Fréquence Lexique mot2	450
Tableau 5.8. Liaisons obligatoires anglophones. Fréquence Lexique mot2	451
Tableau 5.9. Liaisons obligatoires, francophones. Fréquence Lexique cooccurrence	452
Tableau 5.10. Liaisons obligatoires, anglophones. Fréquence Lexique cooccurrence	452
Tableau 5.11. Facteurs linguistiques et liaisons obligatoires pour les trois groupes	453
Tableau 5.12. LO, facteurs linguistiques et individuels/sociaux chez les apprenants	455
Tableau 5.13. Liaisons obligatoires, tous apprenants. Trois Temps.	456
Tableau 5.14. Liaisons obligatoires, apprenants, Groupe 1. Trois Temps.	456
Tableau 5.15. Liaisons obligatoires, apprenants, Groupe 2. Trois Temps.	457
Tableau 5.16. Liaisons obligatoires, apprenants, Groupe 3. Trois Temps.	457
Tableau 5.17. Liaisons obligatoires, apprenants, Groupe 4. Trois Temps.	458

## Liste des figures et graphiques

Figure 1. Le rôle de la prononciation dans Dufeu (2008 : 4)	12
Graphique 4.1. Production de liaisons et enchaînements tous francophones	351
Graphique 4.2. Productions de toutes liaisons et enchaînements, tous les francophones	353
Graphique 4.3. Facteurs participant à l'acquisition des liaisons	384

## Liste des Annexes

Annexe 1 : les directives du Ministère de l'Éducation de l'Ontario (prononciation)	502
Annexe 2 : description du Cadre Européen Commun de Référence (CECR)	508
Annexe 3 : questionnaire 1 distribué aux étudiants dans la session 1	509
Annexe 4 : questionnaire sur ce que les étudiants ont pensé de la leçon de phonétique	513
Annexe 5 : questionnaire pour les francophones	514
Annexe 6 : les photocopies pour le cours de phonétique	515
Annexe 7 : texte de l'étude avec codage des contextes liants	521
Annexe 8 : Les contextes syntaxiques du texte à lire	522
Annexe 9 : Fréquence du mot1 et du mot2 dans <i>Lexique</i>	525
Annexe 10 : Fréquence du mot1 et du mot2 dans <i>Vocabprofil</i>	528
Annexe 11 : Fréquence mots1 et mots2 avec <i>Lexique</i> et <i>Vocabprofil</i>	531

## Préface

Plusieurs événements ont contribué au choix du sujet de cette thèse : *l'acquisition de la liaison et de l'enchaînement en français L2: le rôle de la fréquence.*

Un ami ukrainien qui parlait parfaitement le français m'a raconté l'histoire suivante. Il avait passé un entretien pour un poste de professeur de français langue étrangère (FLE) dans une école de langue en France. L'une des épreuves consistait à enseigner un cours de langue à des apprenants du français Langue Seconde (L2) de niveau avancé. L'un des apprenants lui a posé la question suivante : « Doit-on faire la liaison avec la conjonction *mais* ». Mon ami n'a pas su répondre à cette question et ... n'a pas obtenu le poste. Quand il m'a posé la question, il y a une quinzaine d'années, je n'ai pas su y répondre bien qu'étant francophone et ayant des années d'expérience en enseignement du FLE. Cela m'a intriguée et m'a poussée à me documenter sur la question.

Des années plus tard, en enseignant le français aux étudiants canadiens anglophones, je me suis trouvée face à plusieurs questions auxquelles, une fois de plus, je n'ai pas pu répondre.

1. La règle est pourtant simple ! Il suffit de prononcer la consonne finale du mot1 avec le mot2. Je leur ai dit et expliqué déjà plusieurs fois, alors pourquoi les apprenants ne font-ils pas les liaisons ?
2. Après avoir expliqué les règles des liaisons plusieurs fois, après avoir corrigé les erreurs des apprenants plusieurs fois, j'ai constaté que, soit ils ne faisaient pas les liaisons, soit ils continuaient à prononcer la consonne finale séparément du mot2 et à dire par exemple /ilz ariv/, au lieu d'enchaîner la consonne au mot suivant.

Comprendre une règle n'est-il pas suffisant pour appliquer une règle « simple » ?

3. Si expliquer une règle, répéter la règle et corriger les erreurs ne semble pas aider les apprenants, alors que peuvent faire les enseignants pour aider leurs étudiants à améliorer leur prononciation ?

Enfin, lorsque mon fils qui apprend le français comme L2 dans une école francophone a commencé à apprendre à lire et à écrire en première année, j'ai remarqué un phénomène similaire à ce que j'avais remarqué chez mes étudiants adultes anglophones. J'ai eu beau lui lire de petites histoires en français, le faire lire presque tous les jours, lui expliquer qu'il faut prononcer *les animaux* /le zanimo/, il continuait à prononcer les mots qu'il lisait séparément /le animo/ alors qu'il faisait pourtant la liaison dans /le zami/. Je commençais à me décourager quand pratiquement du jour au lendemain, il s'est mis à faire les liaisons, même avec des séquences relativement rares qu'il n'avait probablement jamais entendues. C'était comme si un cap était passé, comme s'il avait franchi un seuil après lequel il commençait à maîtriser le système.

C'est dans le but d'améliorer mes compétences en enseignement, notamment en enseignement de la prononciation du français et de mieux comprendre ces phénomènes d'acquisition que j'ai souhaité trouver des réponses à ces questions. Je remercie donc mes étudiants qui lors de leurs questions, de leurs efforts et de leur volonté d'améliorer leur français m'ont inspirée à trouver de meilleurs moyens de les aider et mes enfants qui ont toujours été une source d'inspiration.



# **INTRODUCTION**

Contrairement à l'anglais, le français est une langue à syllabation ouverte qui favorise les contextes CV (Delattre, 1946). Les enchaînements et les liaisons rajoutent à ce principe en créant encore plus de syllabes ouvertes.

**La liaison** est le phénomène par lequel une consonne latente<sup>1</sup> est prononcée et enchaînée<sup>2</sup> au mot suivant, si celui-ci commence par une voyelle ou un H muet.

Exemples : *Les enfants* se prononce [lezãfã].

*Il est arrivé* se prononce [iletaRive].

**L'enchaînement** est le phénomène par lequel la consonne finale fixe<sup>3</sup> est resyllabée avec le mot suivant si celui-ci commence par une voyelle ou un H muet.

Exemples : *Une grande enfant* se prononce [yn gRãdãfã].

*Une belle histoire* se prononce [yn belistwaR].

Lorsqu'une liaison a lieu, elle se fait entre deux mots : le premier mot se nomme *mot1* et le deuxième mot auquel la consonne finale du premier mot est attachée se nomme *mot2*.

Exemple : *les* (mot1) \_ (z) *enfants* (mot2)

Selon Dubois et al. (2007 : 342) :

En linguistique traditionnelle, le *mot* est un élément linguistique significatif composé d'un ou de plusieurs phonèmes ; cette séquence est susceptible d'une transcription écrite [...] comprise entre deux blancs ; dans ses divers emplois syntaxiques, elle garde sa forme, soit totalement, soit partiellement (dans le cas de la flexion). Sur le plan sémantique, le mot dénote un objet (substantif), une action ou un état (verbe), une qualité (adjectif), une relation (préposition), etc. C'est cette définition qui est retenue en lexicographie. Une telle définition rencontre diverses difficultés portant : [...] d) et, surtout sur le fait que cette définition ne peut rendre compte des expressions figées ou des mots composés de plusieurs unités significatives qui ne sont pas réunies par des traits d'union dans l'écriture. [...]

---

<sup>1</sup> Une consonne latente est non prononcée si suivie d'une consonne

<sup>2</sup> Sauf cas particuliers dont parle Encrevé (1988)

<sup>3</sup> Une consonne fixe est toujours prononcée, que le mot suivant commence par une voyelle ou une consonne

D'après cette définition, si deux mots sont unis par un trait d'union, il ne s'agit que d'un seul mot. Par contre, si les mots sont séparés par des blancs, ce sont des mots différents. Si on considère « le mot » *pomme de terre* d'un point de vue sémantique, il s'agit d'une seule unité, donc d'un seul mot, alors que selon la définition, il s'agit de plusieurs mots séparés par des blancs. Autrement dit, la notion de *mot* est une notion écrite qui dépend des blancs et des traits d'union qui séparent ou unissent les mots.

Mais si seuls les blancs et les traits d'union déterminent qu'il s'agit d'un ou plusieurs mots, que penser d'une séquence telle que *vient-il* ? Celle-ci comprend un trait d'union, et s'agit-il d'un seul mot ? De plus, nous voyons avec le mot *Moyen-âge*, que le trait d'union n'est pas une notion fiable, puisque ce mot peut s'écrire avec et sans trait d'union (Petit Robert, 2002 : 1688).

En outre, d'un point de vue lexical et phonétique, si on analyse, par exemple, le mot composé *sous-entendu*, on ne peut pas considérer qu'il se compose de deux sèmes (non susceptibles de réalisation indépendante), ni même de deux lexèmes. Puisque *sous* et *entendu* peuvent être utilisés de façon indépendante, on peut considérer qu'il s'agit de deux *mots*. En fait, un mot composé est un mot qui comprend deux ou plusieurs mots, tout comme l'est une expression figée, qu'il y ait un trait d'union ou non. S'il s'agit de *deux mots*, le mot1 et le mot2 sont susceptibles de lier, on peut alors considérer qu'il y a une liaison dans *sous-entendu*.

C'est ce qu'en ont également déduit les phonéticiens qui s'intéressent aux classifications des liaisons. Delattre (1947 : 157) propose dans sa liste d'expressions figées (contenant une liaison obligatoire) de nombreux mots composés : *États-Unis*, *guet-apens*, *vis-à-vis*, *sous-entendu*, *sous-officier*, etc. Léon (1978 : 126) donne également comme exemples d'expressions figées qui ont une liaison obligatoire de nombreux mots composés, les mêmes que ceux de Delattre (1947) : *sous-officier*, *sous-entendu*, *États-Unis*, *Champs-Élysées*, *Ponts-*

*et-Chaussées*, etc. De même pour Valdman (1993 : 159), Charliac et Motron (1998 : 16), etc.

La liaison s'étudie généralement avec l'enchaînement, puisque les liaisons sont le plus souvent enchaînées. Selon les définitions de Fougeron et Delais-Roussarie (2004) :

L'enchaînement (€) est un phénomène de sandhi externe par lequel des mots adjacents se trouvent liés les uns aux autres dans la chaîne parlée. Dans une description traditionnelle, l'enchaînement induit une resyllabation de la consonne finale d'un mot avec la voyelle initiale du mot suivant («une grande amie»: [yn.grã.da.mi]). Ainsi la structure syllabique des formes lexicales isolées ([yn] + [grãd] + [ami]) est réorganisée dans la chaîne parlée au-delà des frontières de mots créant un non-alignement entre frontières syllabiques et lexicales. La liaison (£) est un cas particulier d'enchaînement, spécifique au français, dans lequel une consonne latente à la fin de certains mots est réalisée seulement si le mot suivant commence par une voyelle («un petit ami» [ẽ.pøti.ta.mi]). Les liaisons sont généralement enchaînées, la consonne latente étant resyllabée avec la voyelle du mot suivant. (Fougeron et Delais-Roussarie, 2004 :1)

Il existe aussi d'autres différences entre la liaison et l'enchaînement :

- Les consonnes d'enchaînement ne changent pas de prononciation, et se prononcent telles qu'elles sont écrites. Les consonnes de liaison peuvent changer de prononciation : le x ou le s se prononcent /z/ et le d se prononce /t/ (Pernot, 1937 : 336).

Exemple : *sens idéal* [s] en enchaînement vs *sans idéal* [z] en liaison<sup>4</sup>.

- Les consonnes de liaison se limitent à /z/, /t/, /n/, /r/, /p/ (et /v/ pour *neuf*) tandis que les consonnes d'enchaînement comprennent presque toutes les consonnes.
- Normalement, la liaison se fait à l'intérieur de groupes syntaxiques et dans des conditions particulières, alors que les enchaînements sont réalisés, dans un tempo normal ou rapide, dans tous les contextes, mais dans un tempo plus lent, ils ne sont réalisés qu'à l'intérieur de groupes rythmiques (Laeufer, 1994 : 23). Selon les termes de Laeufer (1994 : 3) « liaison is thus best analyzed as syllabification, enchaînement as resyllabification ».
- Comme le précisent Charliac et Motron (1998 : 15) : « à la différence de la liaison [...], l'enchaînement consonantique est toujours obligatoire ». En d'autres termes, contrairement à la liaison, il n'y a pas d'enchaînements obligatoires, facultatifs ou interdits.

---

<sup>4</sup> Exemple donné dans William (1991)

- Il existe aussi des différences acoustiques entre les consonnes de liaison et consonnes d'enchaînement, différences non perceptibles à l'oreille (Spinelli et Meunier, 2005).

En somme, les liaisons et enchaînements ont des règles d'utilisation différentes qui compliquent encore leur traitement (en particulier pour les apprenants du français L2).

De plus, la tendance à éviter l'hiatus explique des phénomènes tels que l'élision, la consonne épenthétique en français standard (*va-t-il*), l'ajout de *l' + on* après *si* ou *et* (*si l'on considère, et l'on considère...*), les changements de prépositions (*en Avignon*, au lieu de *à Avignon, en Iran vs au Canada*), et les adjectifs masculins variables tels que *ce-cet, beau-bel*, etc. La contrainte anti-hiatus se manifeste également à l'oral familier, que ce soit avec certaines élisions (*t'as* pour *tu as*), avec la consonne épenthétique (*donne-moi-z-en*), ou avec des *z* tels que *quatre-z-yeux*.

Cela ne veut pas dire que les francophones ne font jamais d'hiatus et que s'ils ont le choix, ils choisiront systématiquement de ne pas faire l'hiatus. Si c'était le cas, toutes les liaisons seraient obligatoires et on ne trouverait aucun hiatus. Or, on en trouve à l'intérieur de mots (*aérer, éolienne, oasis*) comme en frontières de mots (*est arrivé, et il*).

Toutefois les cas d'hiatus sont beaucoup plus restreints que les cas d'évitement d'hiatus. D'une part, ils sont rares à l'intérieur des mots, d'autre part, après /i/ et /u/ ils deviennent des semi-consonnes, (*pluie, fouette*). Enfin, certains mots ayant un hiatus sont des mots d'origine étrangère (*ouaouaron*, d'origine amérindienne ou *oasis* d'origine égyptienne), certains mots étaient au départ une onomatopée (*ouistiti*) et d'autres proviennent de l'ajout d'un préfixe à un mot commençant par une voyelle (*réagir*).

De plus, lorsque certaines liaisons sont interdites, il y a des raisons particulières telles que celles visant à éviter certaines ambiguïtés dans des cas d'homophonie en plus de raisons étymologiques (mots d'origine étrangère).

Exemples : *Il est beau et intelligent. Vs Il est intelligent.*  
*Ce sont des héros. Vs Ce sont des zéros.*

En fait, les cas d'hiatus demanderaient une étude particulière. Par exemple, il est possible que les structures syntaxiques soient en partie liées à certaines contraintes phonétiques. Alors que certains mots à initiale consonantique se placent dans une position, des mots à initiale vocalique de même catégorie syntaxique peuvent difficilement se placer dans cette même position.

Cela pourrait expliquerait par exemple la position des adjectifs qui sont majoritairement placés après le nom. Certains adjectifs peuvent se placer avant le nom, seulement au féminin : « *une blonde enfant* » ou devant une consonne « *un grossier personnage* » pour éviter une liaison « désagréable » ou impossible : « *un blond t-enfant, un grossier r-animal* ». Certaines séquences peuvent se trouver à l'écrit (*froid hiver*), tandis qu'à l'oral on changerait l'ordre des mots pour éviter la liaison impossible et l'hiatus, en disant plutôt : un *hiver froid* (Morin, 2005).

La langue française comprend de très nombreuses règles pour éviter l'hiatus (même si celui-ci existe quand même), dont les liaisons et enchaînements. Cela ne veut pas dire que la contrainte anti-hiatus puisse, à elle seule, justifier l'existence et les productions de liaisons et d'enchaînements. Par contre, on peut dire que c'est un des facteurs qui contribuent à ce phénomène.

Les francophones semblent « naturellement » avoir tendance à « lier » les mots pour éviter l'hiatus, ce qu'on voit dans les études chez les enfants qui font très tôt des liaisons, rajoutent des consonnes d'épenthèse ou font de fausses liaisons, mais évitent généralement l'hiatus (Chevrot, Dugua et Fayol, 2005 ; Dugua, 2005 et 2006 ; Chevrot, Chabanal et Dugua, 2007). La contrainte anti-hiatus et la syllabation ouverte expliquent en partie la liaison et l'enchaînement consonantique et vocalique.

Parmi les liaisons, on compte les liaisons obligatoires, facultatives et interdites. L'utilisation des termes *obligatoires*, *facultatives* et *interdites* n'a pas ici de connotation

particulière (prescriptive ou autre). Nous entendons par *obligatoires* (catégoriques) celles qui sont réalisées à plus de 95%, c'est-à-dire dans la grande majorité des cas, par *interdites* celles qui sont réalisées à moins de 10%, c'est-à-dire celles qui sont la plupart du temps non-réalisées, et par liaisons *facultatives* (optionnelles ou variables) celles qui ne sont ni obligatoires, ni interdites.

Nous utilisons les termes *obligatoires*, *facultatives* et *interdites*, parce que ce sont les termes qui semblent être les plus souvent utilisés dans les manuels de phonétique, par l'Académie Française et par la majorité des chercheurs dont, entre autres, Delattre (1955, 1956), Encrevé (1988), Boula de Mareüil et Adda-Decker (2002), Dugua (2005), Pagliano et Laks (2008), dans le projet PFC, etc.

Les liaisons représentent un sujet d'étude intéressant pour plusieurs raisons. Tout d'abord, elles sont uniques au français, sauf le cas particulier, tels que la liaison en anglais britannique avec le /r/ que mentionnent Lodge, Armstrong, Ellis, et Shelton (1997: 96) : « [...] liaison is also found in English, which in its standard British variety has the liaison consonant /r/ in a word like *hear*, which is silent before a consonant or pause, but pronounced before a vowel. Thus in the sequences: 'I can't hear a noise' /r/ reappears. »

Les liaisons sont une survivance du passé. Les consonnes finales étaient toutes prononcées jusqu'au XII<sup>e</sup> siècle et elles ont progressivement cessé de l'être. C'est l'usage qui a déterminé que seules les liaisons dans les séquences les plus fréquemment utilisées sont prononcées de manière systématique (avec les liaisons obligatoires), ou de façon variable dépendant de certaines conditions (les liaisons facultatives). Les liaisons interdites se trouvent généralement dans des catégories de mots rares d'origine étrangère (autre que grecque ou latine), commençant par une semi-consonne (y) ou un h aspiré, ou dans des séquences pouvant créer des confusions ayant des homophones.

Les francophones majoritaires maîtrisent tous les liaisons les plus fréquentes. La

maîtrise des liaisons interdites est moins uniforme et la connaissance de quand et comment utiliser les liaisons facultatives (avec la bonne consonne de liaison) et la capacité d'adapter son registre de langue (y compris la production de liaisons facultatives plus rares) dépend grandement de l'éducation et du milieu socioéconomique des locuteurs.

La liaison suscite l'intérêt des linguistes depuis quelques décennies. Des théories ont été élaborées afin d'expliquer pourquoi on fait des liaisons, mais aussi et surtout afin de trouver des règles qui régissent un principe universel expliquant les cas de liaisons, de non-liaisons et d'erreurs de liaisons pour le français langue maternelle (L1). Les grammairiens, phonéticiens et enseignants du français langue seconde ou étrangère (L2) ont tenté de décrire les phénomènes, de proposer des classifications, et de fournir des listes à l'usage des apprenants du français L2. Toutefois, les catégorisations classiques de liaisons obligatoires, facultatives et interdites sont pour le moins problématiques pour plusieurs raisons.

D'une part, les liaisons obligatoires ne sont pas toutes réalisées à 100%, et certaines sont plus réalisées que d'autres. Parmi les liaisons interdites, certaines sont réalisées. Et parmi les liaisons facultatives, certaines sont très fréquentes, certaines sont moyennement fréquentes, d'autres sont rares ou très rares.

D'autre part, les règles syntaxiques ne justifient pas pourquoi, à l'intérieur d'une même catégorie syntaxique, certaines liaisons se font systématiquement (par exemple avec *gros*, *petit* ou *grand*), tandis que d'autres se font beaucoup moins (*léger*, *premier*). En fait, aucune théorie (à ma connaissance) ne permet de prédire si des liaisons vont être produites ou non, mis à part le contexte [dét. + nom] ou [pronom clitique + verbe ou pronom] (*nous\_avons*, *il les\_aime*, *allons nous-en*, etc.).

Aucune théorie ne peut répondre de façon définitive à certaines questions. Pourquoi dans un même contexte phonétique et syntaxique, certaines liaisons sont-elles réalisées et d'autres non (voir Sampson, 2001) ? Pourquoi il y a de la variation qui dépend des contextes



phonétiques et syntaxiques, mais aussi des locuteurs eux-mêmes (Ågren, 1973 ; Armstrong, 2001; Ashby, 2003) ? Pourquoi les enfants et les adultes L1 font-ils certaines erreurs (Morin, 2005) ? Pourquoi trouve-t-on des cas de liaisons lorsque le mot2 commence par une consonne ? (On prononce la consonne finale muette de *vingt* dans 22, 23, etc. et on peut prononcer la consonne finale de *quand* même s'il est suivi d'un mot à initiale consonantique.).

Les liaisons facultatives, et dans une moindre mesure les liaisons obligatoires et interdites, constituent chez les francophones un marqueur sociolinguistique. Leur réalisation dépend notamment de l'âge, du sexe, du milieu socioculturel et du niveau d'éducation du locuteur, mais aussi de la situation de communication et du type d'énonciation. Pour plus de détail sur la variation dans les liaisons, voir par exemple, Ågren (1973), Malécot (1975), Gadet (1989) ou Ashby (1981, 1988, 2003).

Les francophones ont une certaine intuition (dont l'appréciation dépend notamment du niveau d'instruction) de ce qui est juste ou faux, de ce qui est bizarre et/ou drôle. Ils jugent ainsi un locuteur de façon inconsciente d'après sa prononciation, dont les liaisons. On voit cela par exemple dans l'étude de Léon et Tennant (1990) et dans des anecdotes, imitations et sketches de comédiens et imitateurs connus.

Le non-enchaînement des liaisons comme le font les hommes politiques (Encrevé, 1988) est perçu pour bon nombre de francophones comme ridicule et risible. Passy (1892) [Cité dans Durand et Lyche (2008 : 49)], semble déjà parler de la liaison non-enchaînée, qui selon lui est ridicule, donc drôle, dans le passage suivant :

Et de fait, on cesse les liaisons dès qu'il y a arrêt. Rien de plus risible qu'une liaison faite mal a (sic) propos. C'est une idée prononcé (sæt, ynide), fait croire qu'on a le hoquet. Un professeur prononçait des phrases comme *la première est excessivement facile* en s'interrompant après est ; (la prəmʃɛr ɛt ɛksesivmã fasil). La première fois que nous l'avons entendu, ça a été un éclat de rire général. »  
(Passy, 1892 : 42)

La fameuse histoire de Nyrop (1935), citée dans Delattre (1955 : 44) dépeint les ennuis que peuvent éprouver les Français à ajuster leurs liaisons à une situation précise :

"Un jour, dans une pièce de Mme de Girardin, La Joie fait peur, la jeune actrice chargée du rôle de l'ingénue dit, en parlant de fleurs qu'elle avait plantées avec son frère: *nous les avions plantées s-ensemble*, en faisant sentir l's. Mme de Girardin bondit sur sa chaise: Pas d's! Pas d's! s'écria-t-elle. Planté ensemble. Vous n'avez pas le droit de faire de pareilles liaisons à votre âge! Je me moque de la grammaire! Il n'y a qu'une règle pour les ingénues, c'est d'être ingénues! Cet affreux s vous vieillirait de dix ans!" (Nyrop, 1935 : 136)

Dans le sketch humoristique des Nuls, *Les liaisons dangereuses*<sup>5</sup> qui reflète les difficultés phonétiques et morphosyntaxiques que rencontrent les acteurs, les erreurs sont basées sur ce qu'on entend dans les médias, dans les films, principalement avec les liaisons facultatives et interdites, et sont en cela représentatives des erreurs de liaisons réellement commises.

Pour des francophones qui ne connaissent ni les règles phonétiques de façon formelle, ni les termes linguistiques qui les décrivent, les erreurs de prononciation sont perçues comme étant des « fautes de grammaire » ou des « fautes d'accord », comme le dit le présentateur de la saynète. La perception qu'a un locuteur natif de la qualité de la langue parlée et même de la personnalité du locuteur dépend en grande partie de la prononciation qu'a celui-ci. (Tardif et d'Anglejan, 1981).

Les non-natifs ont tendance à « séparer » les mots, et à produire moins de liaisons et enchaînements que les francophones. Pour les apprenants du français L2 « L'habitude de découper les mots est difficile à vaincre, surtout après l'exposition à l'écrit orthographique. Il est donc souvent nécessaire de travailler l'enchaînement des mots comme une difficulté à part » (Lauret, 2007 : 59). D'ailleurs, les phonéticiens sont nombreux à dire que l'égalité syllabique, la syllabation (et resyllabation) du français et en particulier l'enchaînement consonantique sont difficiles à maîtriser pour tous les apprenants du français langue L2 (voir par exemple, Charliac et Motron, 1998 : 7-9).

En plus des enchaînements consonantiques et vocaliques, s'ajoute la difficulté des liaisons. Il faut non seulement savoir quand et comment lier des mots, il faut aussi connaître

---

<sup>5</sup> [http://www.dailymotion.com/video/xtplh\\_les-liaisons-dangereuses\\_fun](http://www.dailymotion.com/video/xtplh_les-liaisons-dangereuses_fun)

les règles de prononciation des consonnes de liaison, les règles de dénasalisation (*moyen-âge*) de voisement (*neuf heures, dix ans*), de dévoisement (*grand immeuble*), ainsi que la hiérarchie des liaisons les plus fréquentes pour savoir quelles liaisons réaliser en priorité et comment les réaliser. Maîtriser cette composante phonétique est loin d'être chose aisée, même pour les francophones pour lesquels cela prend plusieurs années. Les règles sont nombreuses et compliquées, d'où les erreurs des enfants et des adultes L1 (Pallaud et Savelli, 2001). C'est pourquoi, selon Thomas (1998 : 544), la liaison est « rarement enseignée à cause de sa complexité ».

Les liaisons et enchaînements sont considérés dans les manuels de phonétique comme étant une « catégorie à part » des éléments segmentaux (prononciation de sons isolés) et des éléments suprasegmentaux (prosodie, intonation, accentuation, rythme, etc.). Toutefois, les liaisons et enchaînements sont liés aux éléments segmentaux et suprasegmentaux. En effet, la production de liaison dépend de plusieurs facteurs tels que les effacements de schwa, l'accentuation de syllabe, et de nombreux autres facteurs eux-mêmes considérés comme suprasegmentaux et constituent un élément phonétique d'une grande complexité. Par exemple, l'accent d'insistance en début de syllabe du mot<sup>2</sup> peut engendrer un coup de glotte qui empêche la liaison. L'effacement du schwa peut entraîner une assimilation qui peut faire fusionner certains mots et par là-même empêcher la liaison avec les mots fusionnés comme dans *je suis allé*, prononcé [ʒʁi ale].

Les liaisons font partie d'un grand ensemble phonétique (la prononciation) qui joue un rôle primordial dans l'acquisition d'une langue. Les enfants commencent par en apprendre les éléments suprasegmentaux, ce qui, selon Kjellin (1999), permettrait d'apprendre le reste de la langue. Ils apprennent aussi la prononciation de leur L1 avant d'en apprendre les mots et les structures. Dufeu (2008) résume dans la figure 1 le rôle de la prononciation dans l'acquisition d'une L2.

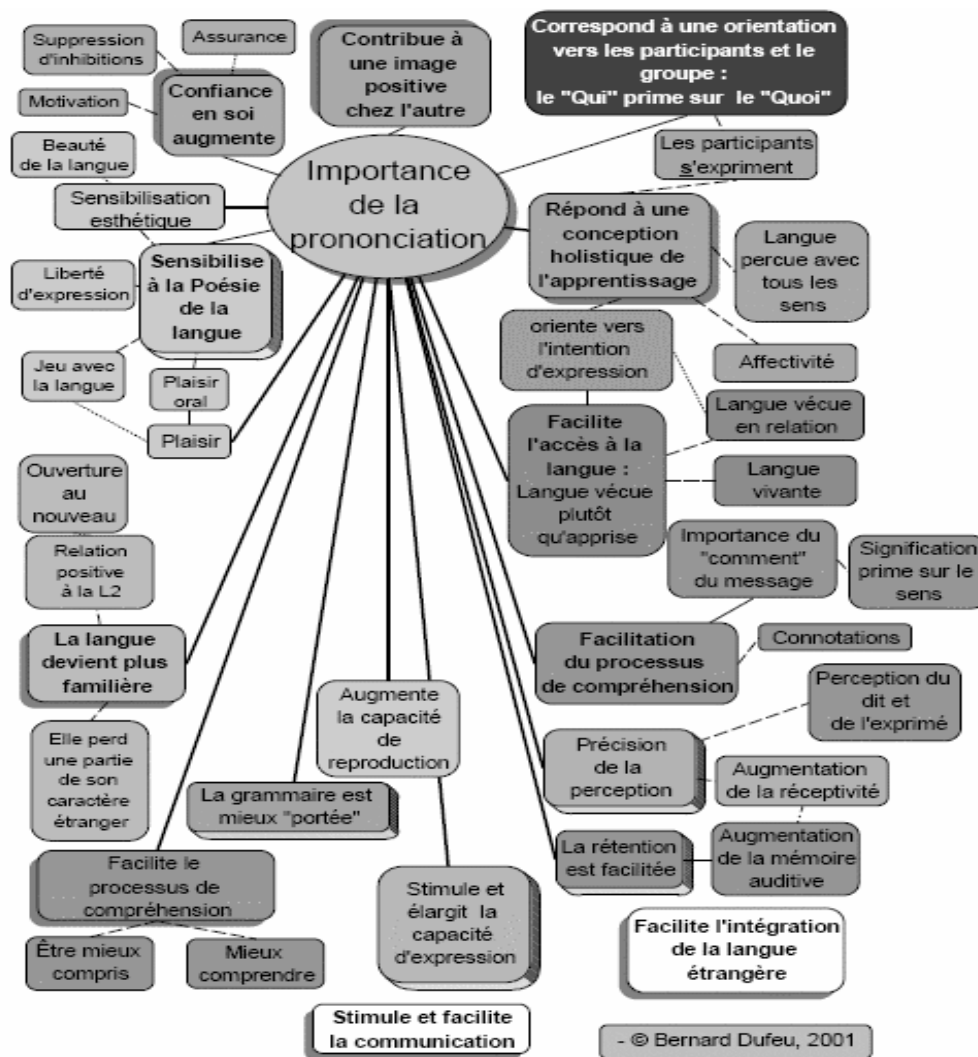


Figure 1. Le rôle de la prononciation dans Dufeu (2008 : 4).  
[http://www.franparler.org/dossiers/phonetique\\_dufeu3.htm](http://www.franparler.org/dossiers/phonetique_dufeu3.htm)

Toutefois, en dépit de son importance, quelles que soient les approches pédagogiques utilisées aujourd'hui, la prononciation a rarement la place qu'elle mérite. Selon Dufeu (2008), cette lacune est due d'une part à l'évolution historique de l'enseignement des langues (grec et latin), d'autre part au fait que la transmission des langues dans les pays occidentaux est principalement écrite.

En outre, les éléments de l'écrit sont mieux connus des professeurs de langue que les éléments de l'oral. Il en résulte que les activités orales, le travail à la maison et les systèmes d'évaluation dépendent souvent d'une source écrite. Si l'oral est évalué, la seule partie évaluée qui est vraiment orale est le plus souvent la fluidité (l'aisance). Sinon, les autres

aspects évalués à l'oral sont les mêmes qu'à l'écrit et les professeurs de langue évaluent surtout la grammaire et le vocabulaire. D'ailleurs, si l'oral n'est pas enseigné en tant que tel (la prononciation, la variation sociolinguistique, les pauses et hésitations, la langue typique de l'oral, etc.), on imagine mal qu'il puisse être évalué en tant que tel.

La prononciation ne fait pas partie des priorités pédagogiques, comme en témoignent, entre autres, les directives du Ministère de l'Éducation de l'Ontario<sup>6</sup> (voir annexe 1). Une bonne prononciation est souvent considérée comme un luxe, comme un travail sur les subtilités de la langue, ce qu'on peut constater notamment dans la description<sup>7</sup> du Cadre Européen Commun de Référence (CECR) (voir annexe 2), où seules l'intelligibilité et la clarté (ce qui est pour le moins subjectif) sont mentionnées, et la seule précision concernant des éléments phonétiques concrets se trouve au niveau C1 (avancé) : « Peut varier l'intonation et placer l'accent phrastique correctement afin d'exprimer de fines nuances de sens ».

Pourtant, il y a ici une inadéquation entre cette composante phonétique et le niveau auquel elle est placée. L'intonation et l'accentuation de syllabe ne constituent pas de « fines nuances », mais sont au contraire le pré-requis pour comprendre et être compris des francophones, c'est-à-dire que ces éléments devraient être appris et requis au niveau A1.

En effet, comment comprendre par exemple la différence entre une phrase affirmative et une phrase interrogative que seule l'intonation différencie ? Si un locuteur anglophone utilise en français l'intonation descendante de l'anglais à chaque fin de groupe rythmique, un francophone non habitué à cet accent pensera que c'est l'intonation descendante marquant la finalité en français, et pensera donc que c'est la fin d'une idée, alors que la phrase continue.

Ainsi, le locuteur francophone sera tellement distrait par l'antinomie de l'intonation descendante et de la continuation inattendue qu'il ne comprendra peut-être pas l'idée de la phrase même s'il en comprend les mots. De plus, une intonation monotone ou descendante, au

---

<sup>6</sup> <http://www.edu.gov.on.ca/eng/curriculum/secondary/>

<sup>7</sup> [http://www.coe.int/t/dg4/linguistic/Source/Framework\\_FR.pdf](http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf) p.92

ton plus bas que le français, couplée à une accentuation de syllabes fautive, risque au pire de ne pas être comprise, au mieux d'être (mal) interprétée comme un signe de tristesse, de déception, de dépression, d'ennui ou d'étrangeté. C'est-à-dire que le message de l'apprenant aura de fortes chances d'être mal interprété.

Si la prononciation n'est ni enseignée, ni évaluée, si elle est absente des manuels et que les enseignants ne sont pas formés pour l'enseigner ; si la priorité est la grammaire, le vocabulaire et la fluidité ; et si l'input reçu se limite à quelques heures par semaine, il paraît impossible que les apprenants maîtrisent la prononciation dont les liaisons et enchaînements font partie. Nous pensons donc que les anglophones ne peuvent pas maîtriser les liaisons et enchaînements comme les francophones parce qu'ils n'en ont pas eu les moyens.

Les études sur les liaisons des apprenants L2 sont rares (Mastromonaco 2000, Thomas 2002, 2004 et Howard 2004, 2005, 2006). Les productions de liaisons des étudiants anglophones comparées aux productions de francophones suggèrent que les étudiants de niveau intermédiaire (B1-B2 d'après le CECR) maîtrisent les liaisons, à un niveau proche des locuteurs natifs. Toutefois les résultats de ces études sont en rapport direct avec ce qu'on cherchait à tester : la maîtrise des liaisons obligatoires les plus fréquentes telles que les liaisons entre un déterminant et un nom ou entre un pronom et un verbe.

Si les sujets des études maîtrisent certains aspects des liaisons obligatoires les plus fréquentes, cela ne veut pas dire qu'ils maîtrisent forcément les liaisons obligatoires aussi bien que des locuteurs natifs étant donné l'input restreint qu'ils ont reçu. En fait, ce qui serait étonnant c'est qu'ils obtiennent des résultats proches des francophones alors qu'ils n'ont pas reçu un input suffisant (sauf s'ils ont été en école francophone) pour apprendre des règles phonétiques et grammaticales naturellement, comme le font les enfants francophones, que le français entendu n'était pas toujours parlé par des francophones, et qu'ils n'ont le plus souvent pas eu un entraînement suffisant notamment en orthoépie.

De nombreux aspects des liaisons et des enchaînements n'ont pas ou peu été étudiés, en particulier la production d'enchaînements en L2.

Notre étude a plusieurs buts. D'une part évaluer les connaissances des étudiants anglophones sur les liaisons et les enchaînements, comparées aux productions de francophones lisant le même texte. D'autre part, tester le rôle de différents éléments de l'enseignement de la prononciation (ici les liaisons) sur les progrès des étudiants, et déterminer les facteurs qui semblent jouer le rôle le plus important dans l'acquisition des liaisons. Enfin, cette étude a pour but général de mieux comprendre le processus d'acquisition des liaisons pour pouvoir trouver des méthodes mieux adaptées à leur enseignement et de façon plus globale l'enseignement de la prononciation en L2.

Afin de mieux tester les connaissances des étudiants, cette thèse analysera plus en détails la production de liaisons obligatoires et enchaînements des étudiants anglophones, et dans une moindre mesure, les liaisons facultatives et interdites.

Nous comparerons les lectures à haute voix de différents groupes de francophones et d'apprenants anglophones lisant un texte spécialement créé pour déterminer les productions de liaisons d'après certains contextes syntaxiques et phonétiques. Nous examinerons ainsi comment les francophones traitent les informations et quelles sont les différences entre les francophones et les anglophones.

Nous souhaitons étudier les progrès que font quatre groupes d'apprenants majoritairement anglophones après une petite leçon de phonétique portant principalement sur les liaisons obligatoires. Pour cette étude, dans le cadre du Modèle Basé sur l'Usage (MBU), nous aborderons les recherches portant sur les liaisons en L1 et L2. Nous expliquerons d'abord le MBU et le rôle de la fréquence et de la productivité des structures dans l'acquisition de diverses composantes linguistiques.

Les recherches mentionnées portent sur des éléments très différents tels que l'acquisition de structures morphosyntaxiques, lexicales ou phonétiques, en L1 comme en L2, et elles portent aussi sur des groupes linguistiques variés : par exemple les germanophones apprenant l'anglais L2, les suédophones apprenant le français L2, les enfants francophones apprenant le français L1, les enfants anglophones apprenant le français L2 ou les étudiants anglophones adultes apprenant le français L2. Quels que soient les éléments linguistiques étudiés, quelle que soit la langue maternelle et la langue cible des sujets, les résultats vont dans le même sens : la fréquence et la productivité des structures jouent un rôle majeur dans l'acquisition, en L1 comme en L2.

Nous discuterons du rôle de la fréquence sur l'acquisition d'une L1 et d'une L2 en général, et plus précisément sur les liaisons, en abordant leur acquisition chez les enfants L1, les erreurs chez les adultes L1, la fréquence et le texte de notre étude, la procédure de notre étude, les liaisons chez les adultes francophones majoritaires et minoritaires, et les liaisons chez les apprenants du français L2.

Nous discuterons enfin de certains résultats de notre étude et de l'enseignement de la prononciation. La conclusion abordera les implications générales de notre étude sur l'enseignement de la phonétique corrective ainsi que ses limites. Une meilleure compréhension de l'acquisition des liaisons chez les francophones et les difficultés des apprenants L2 pourrait contribuer à développer des approches pédagogiques et à créer du matériel pédagogique visant à y remédier.



# **CHAPITRE 1**

## **LES RECHERCHES SUR LES LIAISONS**

## **ET LE MODÈLE BASÉ SUR L'USAGE (MBU)**

## Introduction

De nombreux linguistes ont tenté d'expliquer les liaisons à la lumière des théories en vigueur dans leur discipline, notamment en phonologie (pour un résumé de ces théories, cf. Mastromonaco, 2000 : 26-76) et en syntaxe (Mastromonaco, 2000 : 77-99). Toutefois aucune de ces théories n'est parvenue à donner une réponse définitive aux questions suivantes :

1. Pourquoi certaines liaisons sont-elles obligatoires et certaines facultatives dans exactement le même contexte phonétique et syntaxique ? Par exemple, dans le contexte [adverbe + adj.] (C'est très\_intéressant) la liaison est obligatoire après *très* mais elle est facultative avec *pas* (Ce n'est pas\_intéressant) (voir par exemple Léon, 1978 : 129 ou Abry et Chalaron, 1994 : 110). En effet, dans l'étude de Mallet, (2008 : 251), la liaison est réalisée à 96.55% avec *très*, mais à 1.36% avec *pas*, alors qu'il s'agit d'adverbes monosyllabiques ayant tous deux la consonne de liaison /z/.

2. Pourquoi y a-t-il des exceptions? Pourquoi y a-t-il des liaisons interdites, notamment avec les H aspirés et les semi-consonnes, dans un contexte de liaison obligatoire (dét. + nom). L'étymologie seule ne peut expliquer pourquoi la liaison est obligatoire avec certains h de même origine étymologique alors qu'elle est interdite avec d'autres (Southworth, 1970). Par exemple, le mot *hérisson* est d'origine latine mais a pourtant un h aspiré (donc une liaison interdite avec un déterminant) tandis que le mot *héritage*, venant lui aussi du latin, a un h muet (donc une liaison obligatoire avec un déterminant).

3. Pourquoi parmi les liaisons obligatoires et facultatives certaines sont produites plus que d'autres ? (voir notamment Mallet, 2008).

4. Pourquoi les enfants francophones passent-ils par trois étapes pour apprendre les liaisons (Dugua, 2005) ?

5. Pourquoi retrouve-t-on souvent un ordre d'acquisition similaire (acquisition des langues en général) chez les enfants L1 et chez les apprenants L2 ? Comme le soulignent Lightbown and Spada (2006 : 82) : « [...] it is more remarkable that we find developmental sequences that are similar in the developing interlanguage of learners from different backgrounds and also similar to those observed in first language acquisition of the same language. »

6. Comment expliquer le fait que les erreurs (« cuirs » et « velours ») que font les enfants francophones sont similaires aux erreurs ou fausses liaisons que font les adultes et se retrouvent dans les mêmes contextes (Desrochers, 1994; Morin, 2005 ; Dugua, 2005), en particulier des liaisons avec formes verbales en liaisons facultatives?

7. Pourquoi dans un même contexte syntaxique et phonétique, la voyelle finale de certains adjectifs antéposés est-elle dénasalisée (*Moyen-âge, bon ami*), tandis que d'autres ne montrent pas une telle dénasalisation (*mon ami*), et avec certains adjectifs se terminant par une voyelle nasale il n'y a pas de liaison du tout (*hautain, malin*) (Sampson, 2001)?

8. Pourquoi certains adjectifs sont-ils prononcés différemment selon le contexte. Par exemple, dans *neuf heures, neuf ans*, le f se prononce /v/ alors que dans tous les autres cas, il se prononce /f/ comme dans *neuf enfants*?

9. Pourquoi fait-on la liaison dans certaines expressions figées et pas dans d'autres ? (Delattre, 1947)

10. Pourquoi certaines personnes n'enchaînent-elles pas les liaisons (Encrevé, 1988 ; Morin, 2005) et comment expliquer que ces cas-là sont les mêmes que ceux pour lesquels il y a des erreurs chez les enfants comme chez les adultes, les cas de liaisons facultatives (Pallaud et Savelli, 2001) ?

11. Comment expliquer la tendance à faire de moins en moins de liaisons, notamment la tendance à faire moins de liaisons facultatives (Ashby, 2003 ; Morin, 2005) ?

Une théorie susceptible d'apporter des réponses à ces questions est la théorie de la fréquence, aussi appelée Théorie Basée sur l'Usage ou Modèle Basé sur l'Usage (dorénavant MBU). Pour mieux comprendre en quoi cette théorie s'applique aux liaisons, nous ferons d'abord le lien entre le rôle de la fréquence et l'enseignement des langues, phénomène connu depuis longtemps des pédagogues qui a déjà sa place dans les approches pédagogiques. Puis nous ferons un survol des recherches sur le MBU en rapport avec l'acquisition d'une L1 et d'une L2. Ensuite, nous analyserons les liaisons et l'évolution des langues, les explications et catégorisations « classiques » et quelques études majeures sur les liaisons.

Enfin, nous verrons plus précisément le lien entre le MBU et la production de liaisons, le MBU et les fausses liaisons (cuirs et velours), ce qui nous permettra de répondre à quelques-unes des questions ci-dessus.

# **1. Recherches sur l'enseignement des langues, le rôle de la fréquence et le MBU**

Avant de nous pencher plus précisément sur les liaisons, considérons d'abord le rôle de la fréquence et de l'input dans l'enseignement des langues secondes.

## **1.1. Survol des méthodes d'enseignement des langues**

Le rôle de la fréquence de l'input dans l'acquisition d'une L1 et d'une L2 est reconnu depuis longtemps. Il suffit de considérer les différentes méthodes d'enseignement ainsi que les recherches en acquisition des langues pour voir que l'input, le rôle de la fréquence d'écoute et de production ont toujours été importants.

Rodriguez Seara (2001) fait un survol de l'histoire des principales méthodes ou approches pédagogiques utilisées dans l'enseignement des langues secondes/étrangères, abordant notamment les recherches de Puren (1988) et Germain (1993).

Lors de l'enseignement de langues mortes, avec la méthode « grammaire-traduction », les apprenants se trouvent en contact avec la langue cible (le grec ou le latin), à l'aide de lectures intensives. Au début, il s'agit de lire et traduire des petits textes ou des extraits, puis de lire des livres entiers dans la langue cible. Le latin étant une langue à déclinaison, il n'est pas possible de la maîtriser sans avoir des connaissances en grammaire. Avec cette méthode il s'agit non seulement de comprendre et de savoir écrire la langue, mais aussi de connaître la culture reliée à la langue. Comme il s'agit plutôt d'accumuler des connaissances théoriques et de maîtriser l'écrit, en compréhension comme en production, cette méthode répond à ces besoins.

Par contre, pour l'apprentissage de langues vivantes, les besoins sont différents : il faut apprendre à communiquer et à maîtriser l'oral. À la fin du XIX<sup>e</sup> siècle, Gouin (1880) analyse comment les enfants apprennent une L1 et crée la méthode « naturelle », ou « Méthode des

séries ». Gouin remarque que les enfants apprennent leur L1 selon un certain ordre. Donc, si des apprenants veulent apprendre une langue comme le font les enfants, ils doivent écouter la langue cible, apprendre et utiliser les mots selon leur ordre de fréquence. Par contre, les séries de phrases à apprendre, l'ordre chronologique et d'autres éléments de cette méthode la rendent difficile à utiliser et surtout l'apprentissage qui s'ensuit est incomplet, comme le sont les suites de cette méthode. Gouin a été le précurseur de méthodes plus modernes et le premier à primer l'importance de l'oral. Par ailleurs, selon Germain (1993) il est le pionnier de l'immersion.

A la fin du XIX<sup>e</sup> siècle, apparaît la méthode « directe » en France et en Allemagne. Au début du XX<sup>e</sup> siècle, des directives gouvernementales imposent cette méthode dans l'enseignement national en France. Là aussi, cette méthode est basée sur l'observation de l'acquisition d'une L1 chez les enfants. Il s'agit alors d'apprendre la langue cible, par l'oral, sans l'écrit, sans règles de grammaire. L'apprentissage se fait, comme pour les enfants, de manière intuitive et inductive, en écoutant et en utilisant la langue.

L'objectif est de communiquer et de maîtriser l'oral. Pour cela, les professeurs sont de langue maternelle et font répéter des phrases et saynètes à leurs élèves de très nombreuses fois. Seule la langue cible est utilisée en classe. C'est une méthode basée sur l'imitation, avec des répétitions intensives et mécaniques. Les professeurs comme la méthode sont stricts, les erreurs des apprenants inacceptables. L'usage de la langue maternelle des apprenants est interdit ainsi que les explications grammaticales ; l'écrit passe au second plan et il est difficile de trouver des professeurs maîtrisant suffisamment bien la langue orale et la méthode.

Une méthode un peu plus souple se développe entre les années 1920 et 1960 : la méthode « active ». Il s'agit surtout d'utiliser la méthode directe tout en donnant un peu plus de liberté aux professeurs et en leur permettant notamment d'utiliser la langue maternelle des apprenants pour des explications grammaticales.

Lors de la seconde guerre mondiale, l'armée américaine a besoin d'une méthode qui permette aux soldats d'apprendre une langue étrangère rapidement. C'est pourquoi Bloomfield crée la méthode « de l'armée ». Cette méthode n'a duré que deux ans mais elle a contribué à la création, en 1950, de la méthode audio-orale basée sur le distributionalisme, sur les travaux de spécialistes en linguistique appliquée tels que Fries (1945) et sur les recherches des behavioristes. La méthode a évolué avec les travaux de Lado (1957).

Le behaviourisme s'est développé avec les recherches de Watson (1924) et de Skinner (1957). Selon cette théorie, on apprend, ou d'après la distinction de Krashen (1982), on « acquiert » une langue en imitant les locuteurs natifs, en formant des habitudes basées sur des stimuli. Selon les adeptes de ces principes, les enfants imitent et répètent un peu comme des perroquets ; l'input (le stimulus) et la répétition se situent au centre de l'apprentissage.

Avec la méthode audio-orale, basée sur les théories behavioristes, les apprenants se concentrent sur l'oral, écoutent et répètent de très nombreuses fois des groupes de mots ou des phrases afin que la parole devienne un réflexe. Il y a un minimum d'explications et un maximum de répétitions avec des exercices structuraux en laboratoire, afin d'imiter les locuteurs natifs. Le professeur, de langue maternelle, propose un modèle que les élèves imitent et reproduisent.

Les difficultés qu'éprouvent les apprenants à former autre chose que des phrases préfabriquées, leurs difficultés à communiquer en situation réelle de communication et l'ennui et le manque de motivation que provoquent les exercices de « *pattern-drill* » en laboratoire de langue, ainsi que la grammaire générative des années 1960 conduisent à une nouvelle méthode, la méthode audio-visuelle. Nous mentionnons cela car lorsqu'on parle de fréquence dans le cadre du MBU, il ne s'agit pas d'écoute et de répétition passive (comme c'était le cas avec cette méthode) ayant pour seul but d'être capable de répéter, mais d'utiliser la fréquence en début d'acquisition comme outil ouvrant la porte à d'autres activités de communication.

Afin de réagir contre l'essor de la langue anglaise et de promouvoir le français à l'étranger, le Ministère de l'Education Nationale crée une commission dirigée par le linguiste Gougenheim et le pédagogue Rivenc, commission chargée de mettre sur pied le français fondamental, (le français de base), à partir de la fréquence de celui-ci dans la langue parlée.

En 1954, l'équipe du CREDIF met sur pied deux listes du français fondamental. La première liste est constituée des 1475 mots les plus fréquents : le français fondamental premier degré, et la deuxième liste des 1609 autres mots les plus fréquents : le français fondamental deuxième degré. A partir de ce français fondamental, des dialogues sont créés de façon à utiliser ces mots dans des textes les mettant en pratique à l'oral, ces mots étant introduits de façon progressive et selon leur ordre de fréquence.

Au milieu des années 1950, Guberina de l'université de Zagreb crée les fondations de la méthode structuro-globale audiovisuelle, ou méthode SGAV. Cette méthode est possible grâce aux progrès techniques qui rendent les bandes magnétiques et les supports visuels plus accessibles, bien qu'encore très chers.

Le CREDIF utilise les méthodes de Guberina et le français fondamental et Rivenc crée en 1962 la méthode « Voix et images de France » qu'il développe en 1964, méthode audiovisuelle française utilisée dans les années 60-70. Cette méthode évoluera avec *Niveau Seuil* (1976) réunissant les principes des méthodes précédentes, les résultats des recherches de ce moment-là et prenant aussi en considération les besoins de l'apprenant.

Plusieurs facteurs compromettent la méthode audiovisuelle : le coût élevé du matériel audiovisuel (laboratoire), la formation des enseignants, le nombre réduit d'apprenants et le nombre d'heures important. De plus, cette méthode permet aux apprenants de communiquer avec des locuteurs natifs, mais ne leur permet pas de comprendre les locuteurs natifs parlant entre eux. Enfin, de nouvelles théories en linguistique (le fonctionnalisme) et en psychologie (le cognitivisme) cèdent la place à l'approche communicative.



Le but du français fondamental était de déterminer les besoins des apprenants et de répondre à leurs besoins. En utilisant les mots les plus fréquents, les apprenants devaient être capables de communiquer rapidement. Toutefois, dans les années 70, des linguistes et psychologues proposent de placer l'apprenant au cœur de l'apprentissage et de se concentrer sur ses besoins avec plus de précision.

Or, les apprenants ont des besoins professionnels particuliers. Il s'agit alors de créer des programmes qui répondent aux besoins d'un public adulte exigeant. Le français fondamental a donné naissance au français instrumental, et ensuite au français fonctionnel avec le Niveau Seuil. Le but est que les apprenants soient capables de communiquer dans une situation professionnelle précise (utiliser le vocabulaire du tourisme, des affaires, etc.).

Bien que l'approche communicative ait commencé à être utilisée en 1975, cela a pris plusieurs années avant qu'elle soit véritablement utilisée par la majorité des professeurs et dans les manuels de langue. Avec cette approche, le rôle de l'input et notamment de « l'input compréhensible » (Krashen, 1982) est là aussi, important.

Le but de cette approche est que l'apprenant soit capable de communiquer et qu'il apprenne une langue comme un enfant apprend une L1 : en écoutant et en pratiquant. Pour cela, l'apprenant doit se trouver dans des situations communicatives favorisant une communication authentique qui l'incite à utiliser un input riche et varié, ce qui est particulièrement le cas dans les écoles d'immersion. Le but est de communiquer à l'oral, de comprendre et d'être compris, au détriment parfois de la précision de la langue.

La première génération de cette approche rejette les explications grammaticales explicites, favorise l'acquisition « naturelle », le but principal étant la fluidité avant la précision, comme en témoigne par exemple l'explication de Lightbown and Spada (2006) :

[Supporters of communicative language teaching] [...] suggest that errors are a natural and valuable part of the language learning process. Furthermore, they believe that the motivation of learners is often stifled by an insistence on correctness in the earliest stages of second language learning. These opponents of the 'get it right from the beginning' proposal argue that it is better to encourage learners to develop 'fluency' before 'accuracy'. (Lightbown and Spada, 2006 : 140)

Lightbown et Spada (2006 : 156) soulignent aussi les inquiétudes des utilisateurs de cette approche : « Over the years, however, educators and researchers began to express concern about students' failure to achieve high levels of performance in some aspects of French grammar, even after several years of full-day exposure to the second language in these programmes. »

En effet, plusieurs études ont montré les limites de cette approche. Par exemple, dans l'étude de Trahey and White (1993), les jeunes apprenants francophones de l'anglais L2 ont été « bombardés » de structures mettant en relief la place et l'utilisation des adverbes en anglais, grâce à un input enrichi. S'ils étaient capables de reconnaître les phrases correctes, ils n'étaient pas en mesure de reconnaître les phrases incorrectes ayant des structures similaires au français.

Être exposé à la langue permet certainement de faire de grands progrès, mais l'exposition seule (sans explications formelles) est souvent insuffisante pour contrebalancer l'interférence de la langue maternelle. De même, Lyster (1994) a remarqué qu'après huit ans d'immersion, les élèves avaient tendance à utiliser quasi systématiquement le *tu* au détriment du *vous*. Par contre, après une formation axée sur la forme, les élèves ont fait des progrès.

Naît ensuite dans les années 1990, la deuxième génération d'approche communicative, ainsi que de nouvelles directives des institutions, rendant cette méthode plus éclectique. Celle-ci allie le fond et la forme, unit d'autres approches linguistiques plus traditionnelles, notamment les explications grammaticales, donnant aux professeurs un peu plus de liberté et peut-être aussi un enseignement plus équilibré tout en préservant les principes de base de l'approche communicative d'origine (Rodiguez Seara, 2001 : 18-20).

S'ajoutent ensuite les habiletés de l'apprenant à devenir autonome, à s'auto-évaluer et à travailler ses stratégies d'apprentissage, ce qui a conduit à la pédagogie ou approche actionnelle prônée par le Cadre Européen Commun de Référence des Langues à partir des

années 2000. Les apprenants sont « des acteurs sociaux » et ils doivent être capables « d'accomplir des tâches » définies par des compétences et des stratégies à utiliser lors d'actes de parole, dans des contextes sociaux dépassant les activités purement langagières.

Toutes ces approches semblent avoir plusieurs points en commun : elles se basent sur l'acquisition d'une L1 chez les enfants, elles reconnaissent le rôle de l'input, de l'écoute et de la répétition et placent l'oral au centre de l'apprentissage. Toutefois, les réactions négatives face aux méthodes audio-linguales et l'impact majeur qu'ont eu les théories générativistes et la grammaire universelle ont contribué à minimiser le rôle de la fréquence dans l'input, en linguistique théorique et appliquée.

La position innéiste soutient que les enfants

Parmi les théories qui ont influencé le MBU (Modèle Basé sur l'Usage), se trouve la théorie du connexionnisme issue des travaux de Hebb (1949) et de Hayek (1952) qui s'oppose au behaviorisme. Pour les behavioristes, les réactions des êtres humains sont conditionnées par un stimulus, au même titre que d'autres réflexes d'ailleurs présents chez les animaux. Par contre, pour les connexionnistes, ce sont la mémoire et les connections neuronales provenant de l'expérience de l'apprenant qui renforcent ou modifient le processus d'apprentissage.

Pour Hayek (1952), le cerveau humain est un système de classification des informations. Selon lui, l'être humain mémorise et classe des informations selon ses expériences, ce qui va modeler son cerveau. Cela n'est possible que si le cerveau est prêt à recevoir, traiter et classer les informations. Le cerveau est un système en perpétuelle évolution basé sur l'expérience et la personnalité du locuteur.

Ces théories connexionnistes ont été développées notamment avec la linguistique cognitive (Langacker, 1987, 1991, 2008) et la neurolinguistique. En tant que linguiste cognitiviste, Lamb (1999) a travaillé à établir un lien entre le fonctionnement des réseaux neuronaux et la linguistique cognitive, c'est-à-dire à mieux comprendre le fonctionnement du cerveau humain pour mieux en dégager des théories sur le développement du langage. Ses recherches ont contribué au développement de logiciels en intelligence artificielle entre les années 1960 et 1980.

Les théories connexionnistes et cognitivistes ont contribué à développer la théorie basée sur l'usage (cf Barlow et Kemmer, 2000) pour laquelle, comme le fait remarquer Inglis (2000), Langacker (1987) a été le premier à utiliser le terme « basé sur l'usage ».

Selon le MBU, et contrairement aux principes générativistes, non seulement l'input n'est pas « limité », mais il est au contraire très riche et surtout parfaitement adapté au développement langagier de l'enfant. Les enfants apprennent une langue en passant par plusieurs étapes : la première consiste à écouter, observer et stocker un inventaire de mots et

de structures. A cette étape, les enfants ne font que répéter et font peu d'erreurs. C'est d'ailleurs cette étape qui pourrait avoir inspiré les behavioristes, car en effet, à cette étape il s'agit surtout d'habitudes et de répétition.

Puis, une fois qu'ils ont un inventaire suffisamment riche et varié, y compris des phrases avec des parties à « emplacements ouverts », ils construisent des classes et des catégories, le tout basé sur un système probabiliste. C'est à ce moment-là qu'ils font des erreurs de surgénéralisation (voir par exemple, Dugua, 2005) dans la création et l'utilisation de leurs classifications. Enfin, ils parviennent à rectifier les erreurs et à parler comme les adultes de leur groupe linguistique. Gouin (1880) avait déjà remarqué que les enfants apprenaient une langue en passant par des étapes, et selon un certain ordre correspondant à la fréquence de l'input.

Eskildsen (2008) résume les principales théories en rapport avec la Linguistique Basée sur l'Usage (LBU), en se référant aux études de Tummers et al. (2005) et de Langacker (1987) dans les termes suivants :

A core principle uniting these theories is the rejection of the syntax–lexicon dichotomy and an embracing of a 'maximalistic language model knowledge in which abstract grammatical patterns and the lexical instantiations of those patterns are jointly included, and which may consist of many different levels of schematic abstraction'. Language knowledge is seen as a structured inventory of symbolic units, that is form–meaning patterns, of varying complexity (from morphemes to full utterances) and abstractness (from fixed, perhaps idiosyncratic, concrete formulas to abstract schematic templates which in turn sanction the single instantiations). (Eskildsen, 2008 : 335)

Comme le précise Eskildsen (2008 : 335-336), la plus grande contribution du MBU à la linguistique appliquée est les recherches en acquisition d'une L1 chez l'enfant. Selon ces recherches, les enfants apprennent les deux côtés de la grammaire : les plus petits éléments et de plus vastes modèles (Tomasello 2003 : 41), alors qu'ils construisent leur L1 à partir d'exemplaires concrets qu'ils ont rencontrés en situations de communication.

Rejetant la pauvreté du stimulus (Tomasello, 2003; MacWhinney, 2004), le MBU considère l'acquisition d'une langue comme faisant partie de l'expérience de l'apprenant,

basée sur des items et dirigée vers des schémas de plus en plus abstraits, allant de formules préfabriquées à des constructions abstraites. Les enfants apprennent d'abord des mots ou expressions spécifiques et construisent petit à petit des représentations générales.

Les progrès et les représentations abstraites que construisent les enfants dépendent de leur base de données et des items qu'ils ont rencontrés. Au fur et à mesure que leur base de données grandit, leurs constructions et les degrés d'abstraction se modifient dans leur interlangue pour créer des structures plus productives.

Bybee (2001 : 6-8) définit les principes du Modèle Basé sur l'Usage de la façon suivante :

1. L'expérience influence les représentations. Les modèles qui comprennent plus d'items sont plus forts, plus accessibles et plus productifs que ceux qui en comprennent moins.

2. Les représentations mentales des objets linguistiques ont les mêmes propriétés que pour d'autres objets. Le cerveau opère de la même façon dans différents domaines.

3. Les catégorisations sont basées sur l'identité ou la similarité, l'organisation et le stockage d'éléments phonologiques perçus (*phonological percepts*).

4. Les généralisations concernant les formes ne sont pas séparées des représentations stockées des 'formas' mais en émergent directement. Les généralisations sur les formes sont exprimées en termes de relations parmi les formes basées sur des similarités phonétiques et/ou sémantiques.

5. L'organisation lexicale fournit des généralisations et une segmentation à différents niveaux d'abstraction et de généralité.

6. Les connaissances grammaticales constituent des connaissances procédurales. Un locuteur natif peut former une phrase grammaticalement correcte, sans pour autant être capable d'expliquer les règles utilisées ou les propriétés nécessaires pour qu'une phrase soit acceptable.

Les chercheurs qui ont le plus contribué à élaborer cette théorie sont Tomasello (1992, 2000a et 2000b), Ellis (2002), Langacker (1987) et MacWhinney (1988, 2004). Analysons plus en détail les études les plus représentatives dans le domaine avant de nous pencher sur d'autres études.

Tomasello (1992) a remarqué qu'au début, bien que l'anglais de sa fille de 2 ans soit « grammatical », son discours consistait surtout à répéter et reproduire des éléments basés sur certains items, et ce, de façon inégale. Il a remarqué que certains verbes étaient utilisés dans certains contextes et que certaines structures étaient parfois uniques à certains verbes. D'où son hypothèse de « l'îlot verbal » (*verb-island*) qui serait au centre de l'apprentissage.

Il a émis l'hypothèse qu'à un jeune âge, les enfants utilisent pour chaque verbe un « îlot » particulier notamment pour sa valeur sémantique. Tomasello (2000a) souligne que certains marqueurs morphologiques sont utilisés seulement avec certains verbes (négation, interrogation ...) et ce, de manière non généralisable aux autres verbes. Par contre, certains verbes sont utilisés avec plus de marqueurs et de façon quasi systématique.

Tomasello (2000b : 68) précise que « les grammaires des jeunes enfants pourraient se caractériser par un inventaire de constructions formées d'îlots verbaux qui définissent alors les premières catégories syntaxiques comme étant basées sur le lexique » (notre traduction).

Dans l'étude de Tomasello (2000a et 2000b), les 12 enfants anglophones âgés de 2 à 3 ans utilisent leurs verbes dans une seule structure en début d'apprentissage. Ensuite, 92% de leurs énoncés initiaux composés de plusieurs mots sont reliés à l'une des 25 premières structures basées sur le lexique, énoncés différents pour chaque enfant. Cela suggère que leur syntaxe se base sur des items et des expressions individuels, et non sur une grammaire innée s'appliquant de façon identique et de la même façon à toutes les structures similaires.

De même, Pine et Lieven (1997) ont remarqué que lorsque les enfants commencent à utiliser les articles *a* et *the* entre 2 à 3 ans, c'est avec des mots différents de ceux qu'ils

employaient avant d'utiliser ces articles. Cela veut dire que les enfants ne sont pas capables de généraliser l'usage de l'article aux noms qu'ils ont appris avant d'utiliser les articles et qu'ils emploient encore, car les noms appris avant cela sont alors employés dans des structures sans l'article. Autrement dit, à ce stade, les enfants utilisent des groupes de mots, sans être capables de généraliser les règles que l'usage de ces mots exige.

Donc, jusqu'à 3 ans, les enfants utilisent les mots, groupes de mots et petites phrases préfabriqués qu'ils ont le plus entendus. Vers 3-4 ans et après, les enfants deviennent plus productifs. Tomasello (2000b : 157) précise que les mêmes résultats ont été observés dans d'autres langues. Cela indique que le développement langagier des enfants est un processus universel commun à tous les enfants indépendamment de leur langue maternelle.

Pour suivre le développement langagier des enfants et les capacités qu'ils démontrent, diverses méthodes ont été utilisées. Dans l'étude de Tomasello et Brooks (1998), les enfants ont reçu un stimulus leur donnant un verbe nouveau. Puis, les expérimentateurs les ont incités à produire des phrases demandant d'utiliser une structure différente de l'input reçu pour les nouveaux mots. La majorité des enfants de cet âge ne sont pas parvenus à utiliser les structures autrement que dans le contexte dans lequel ils les avaient entendues.

Les expérimentateurs ont enseigné de nouveaux verbes aux enfants de 2-3 ans. Certains verbes ont été présentés dans un contexte où ils étaient transitifs, d'autres verbes ont été présentés dans un contexte où ils étaient intransitifs. Les expérimentateurs ont posé des questions aux enfants les incitant à utiliser des structures demandant un verbe transitif. Les résultats ont révélé que les enfants ne pouvaient utiliser un verbe dans une structure transitive que s'ils l'avaient appris dans un contexte transitif. Par ailleurs, ils pouvaient utiliser un verbe dans une structure intransitive, seulement s'ils l'avaient appris dans un contexte intransitif.

Cette étude suggère d'une part que les enfants sont très sensibles à toutes les structures grammaticales, et qu'ils apprennent avec précision toutes les caractéristiques des mots, mais



aussi qu'à ce stade, ils ne sont pas encore capables de généraliser ce qu'ils ont appris à des items qu'ils n'ont pas rencontrés dans un autre type de structure.

Tomasello (2000a et 2000b) fait référence à une étude similaire, celle d'Akhtar (1999) dans laquelle les expérimentateurs ont testé les réactions d'enfants âgés de 2;8, 3;6, et 4;4 ans, pour déterminer comment ils traitent les informations relatives aux structures de phrases: la structure la plus fréquente en anglais, qui est SVO (sujet verbe objet) et la structure non standard en anglais : VSO (verbe sujet objet). Les expérimentateurs ont utilisé des verbes nouveaux et ont demandé aux enfants d'âges différents d'utiliser les pseudo-mots, pour déterminer si les enfants allaient reproduire l'ordre des mots entendu dans les énoncés, ordre contraire à celui de leur langue maternelle ; ou si au contraire, ils seraient capables d'utiliser les nouveaux mots seulement dans un contexte acceptable d'après les règles de la langue maternelle, mais dans un ordre différent de ce qu'ils avaient entendu dans l'input.

Les enfants ont réagi différemment à des âges différents. Les enfants plus âgés ont « corrigé » l'ordre « anormal » des mots afin que les verbes se trouvent dans la structure normale SVO de l'anglais; les enfants plus jeunes ont beaucoup plus souvent utilisé les structures qui avaient été utilisées dans le stimulus pour ces verbes, même si une telle structure n'apparaît jamais en anglais. Les jeunes enfants ont tout de même hésité entre produire un ordre « incorrect » qu'ils avaient entendu, et produire un ordre correct qu'ils n'avaient pas entendu. Cela montrait qu'ils avaient certaines connaissances concernant l'ordre des mots en anglais pour ressentir que ces structures étaient « bizarres » mais que leurs connaissances étaient encore insuffisantes pour surmonter leurs tendances à répéter ce qu'ils avaient entendu.

Tomasello (2000a : 68) souligne aussi que bien que les enfants aient appris très tôt certaines formules basées sur des items, la plupart des enfants possèdent un certain nombre de représentations linguistiques plus abstraites, et ce, dès un jeune âge. Cela suggère que

l'acquisition des items et l'acquisition de représentations abstraites ne sont pas incompatibles. Les deux catégories semblent se développer simultanément bien que les enfants acquièrent et stockent de nombreux items, en majorité avant les classes et catégories abstraites.

Au début, les enfants répètent ce qu'ils entendent puis ils créent des formes nouvelles : des structures à emplacements ouverts (*slots*) (Tomasello et al., 1997). On ne sait pas exactement comment ils les créent, mais on suppose que c'est en observant la variation du discours de l'adulte et la position d'un énoncé d'après sa fréquence. En général, la fréquence d'occurrence d'une expression ou d'un mot (*token frequency*) dans l'expérience linguistique de l'apprenant a tendance à la valider (*entrench*), permettant à l'utilisateur d'accéder à l'expression dans son intégralité et de l'utiliser avec aisance. Par contre, la fréquence d'un type de structure (*type frequency*) détermine les possibilités créatives ou la productivité de la construction (Bybee, 1985, 1995).

Ensemble, comme l'explique Tomasello (2000a : 72-73), les deux types de fréquences, avec les processus d'apprentissage correspondants de l'enfant pourraient expliquer la façon dont les jeunes enfants apprennent à utiliser des expressions linguistiques particulières, dans des contextes communicatifs particuliers et généralisent ensuite ces expressions à de nouveaux contextes basés sur différents types de variations qu'ils entendent. Cela comprend des variations dans les types d'énoncés, dans les structures à emplacements ouverts et dans tous les types de constituants d'une construction.

Quand les jeunes enfants veulent s'exprimer, s'ils ont une expression toute faite, disponible et facilement accessible, ils l'utilisent. S'ils n'en ont pas, ils récupèrent les schémas et items qu'ils ont maîtrisés jusque là et les « coupent et collent » ensemble pour les besoins de la situation de communication. Ce qui est alors le premier choix dans le processus de création est un schéma d'énoncé qui peut être utilisé pour structurer l'acte de parole

comme un tout, avec d'autres items qui remplissent les structures à emplacements ouverts et s'ajoutent à ces fondations.

Tomasello (2000a :77) précise qu'il est important que lorsque les enfants « coupent et collent » les nouveaux mots et structures, ils coordonnent non seulement les formes linguistiques en question mais aussi les fonctions communicatives conventionnelles de ces formes. Autrement, ce qu'ils produiraient serait incompréhensible. Il est aussi important que les structures linguistiques qui sont coupées et collées dans des actes de communication soient assez variées, et comprennent des mots isolés mais aussi des catégories abstraites, des énoncés ou des schémas de phrases partiellement abstraits.

Selon Tomasello (2000a : 72), il est plus probable que les enfants de tous âges surgénéralisent des formes moins souvent entendues que les formes les plus souvent entendues alors devenues « enracinées » (*entrenched*). Par exemple, il est plus probable que les enfants fassent des erreurs de surgénéralisation avec *I come it* plutôt qu'avec *I arrived it*, dans des études où les expérimentateurs les induisent à produire des formes non grammaticales. Cela suggère que non seulement les enfants répètent ce qu'ils entendent, mais aussi que plus ils entendent une séquence, plus il leur semble que c'est la seule forme acceptable. En outre, ils font des erreurs avec les formes moins fréquentes parce qu'ils généralisent les règles des formes fréquentes, aux formes moins fréquentes.

Les enfants de tous les âges sont plus sensibles à la fréquence des verbes pour décider de l'ordre SVO. Même s'ils ont reçu un input ayant un ordre non conforme à l'anglais, les enfants parviennent très tôt à rétablir l'ordre, en employant le bon ordre des verbes et des pronoms. Les résultats vont dans le sens de l'étude d'Akhtar (1999) qui avait utilisé des pseudo-mots. Ici les verbes existent mais ont des fréquences différentes. Les verbes les plus fréquents sont beaucoup plus utilisés dans l'ordre canonique que les verbes les moins fréquents. Les enfants de 3 ans sont plus « productifs » que ceux de 2 ans.

Les résultats des enfants de 3;9 ans n'ont pas été influencés par la fréquence. Ils ont utilisé la bonne structure syntaxique, indépendamment de sa fréquence. Cela suggère qu'une fois les schémas établis, la fréquence ne joue plus un rôle aussi important et ne change pas leur production (Matthews, Lieven, Theakston et Tomasello, 2005). La fréquence joue donc un rôle crucial en début d'acquisition, mais elle est beaucoup moins importante par la suite.

Tomasello (2000b) conclut que les enfants imitent l'input de leur entourage en répétant des séquences préfabriquées jusqu'à l'âge de 3 ans environ, puis, petit à petit, ils créent des catégories et des schémas abstraits. Cela va à l'encontre de l'approche générativiste selon laquelle l'apprentissage se fait vite et sans efforts parce que les enfants ont des compétences dès leur naissance et n'ont qu'à apprendre à les utiliser; et selon laquelle l'apprentissage ne dépend qu'indirectement de l'input (qui ne sert qu'à déclencher des structures syntaxiques innées) et l'apprentissage est créatif dès le début parce qu'il est généré par une grammaire abstraite.

Les recherches de Tomasello ont été poursuivies et confirmées par d'autres chercheurs tels que Nick Ellis<sup>8</sup>. L'article d'Ellis (2002) explique le rôle de la fréquence dans l'acquisition d'une L1 et d'une L2 dans divers domaines linguistiques.

Selon ce dernier, les théories basées sur l'usage soutiennent que l'acquisition du langage est basée sur des exemplaires (*is exemplar based*). C'est l'apprentissage au coup par coup de plusieurs milliers de constructions et l'abstraction en régularités à partir de leur fréquence. Ce qui détermine la productivité des modèles (*patterns*), c'est la puissante loi de la pratique, la concurrence des indices et signaux, la satisfaction des contraintes, l'apprentissage connexionniste et les effets de la fréquence des items et des structures.

Selon Ellis (2002), les régularités d'une langue se développent selon un système basé sur l'expérience du locuteur, à partir de catégories et de modèles particuliers pour arriver à

---

<sup>8</sup> Nous précisons ici le prénom *Nick* pour éviter les confusions possibles avec d'autres linguistes dont le nom de famille est Ellis, tels que par exemple Rod Ellis spécialiste en acquisition des L2.

des constructions plus abstraites. La fréquence joue un rôle majeur pour expliquer la variation sociolinguistique et les changements linguistiques. Il y a plus de chances qu'un élément soit perçu (remarqué) lorsqu'une sensation lui est fortement associée. Lorsqu'une sensation est associée à plus d'une réalité, des processus inconscients pèsent les probabilités et perçoivent celle qui est la plus probable. Pour que la perception d'une langue soit précise et facile, le locuteur doit avoir acquis toute une gamme d'associations et les probabilités qui en découlent pour chaque élément de l'input de cette langue.

Les théories psycholinguistiques et cognitives sur l'acquisition du langage soutiennent que toutes les unités linguistiques proviennent de l'utilisation de la langue. L'apprentissage d'une langue consiste à apprendre les représentations de ces associations qui reflètent les probabilités d'occurrences de mappages (*mapping*) entre la forme et la fonction. La fréquence est un facteur majeur dans l'acquisition parce que les règles d'une langue, à tous les niveaux d'analyse (de la phonologie, à la syntaxe et au discours) sont des régularités de structures qui émergent de l'analyse de toute la vie de l'apprenant pour en déduire des caractéristiques distributionnelles de l'input langagier.

La compréhension d'une langue est déterminée par l'énorme quantité d'informations statistiques relatives au comportement d'items lexicaux dans cette langue. Le locuteur connaît les fréquences relatives avec lesquelles des verbes apparaissent à différents temps, dans certaines structures (aux voix actives ou passives, dans des structures transitives ou intransitives, le type de sujet ou d'objet que prend le verbe et beaucoup d'autres informations). Ces renseignements sont acquis avec l'expérience et l'input qui mettent en relief ces propriétés distributionnelles.

Comme le précise Ellis (2002 : 145), cela ne nie pas l'importance du rôle de l'attention (Schmidt, 1993) dans l'acquisition d'une structure, ni du rôle de l'instruction explicite. Les études des vingt dernières années mettent en relief l'efficacité de l'instruction d'une L2 et

montrent que l'instruction basée sur la forme a de nombreux résultats positifs sur les points enseignés, que le type d'instruction explicite est plus efficace que l'instruction implicite et que l'efficacité de l'instruction explicite de la L2 a des effets durables.

Comprendre comment le cerveau humain est capable de construire des catégories est l'un des domaines de recherches les plus importants en sciences cognitives. Pour cela il y a eu de nombreuses recherches avec des simulations computationnelles. Cela comprend les modèles connexionnistes, les modèles basés sur des réseaux sémantiques, les modèles des prototypes, des schémas et des exemplaires (*exemplars*).

Ellis (2002 : 150) explique que dans les modèles connexionnistes, chaque répétition renforce les connexions entre des unités de traits pertinentes et entre les catégories et ces unités. En ce qui concerne la reconnaissance des mots dans la chaîne parlée, il semble que nous activions le phonème initial d'un mot pour activer tous les autres mots du lexique qui commencent avec le même phonème. Les mots à haute fréquence ont une plus forte activation que les mots à basse fréquence. Les sujets mettent plus de temps à reconnaître des mots à basse fréquence que leurs voisins à haute fréquence parce qu'il est plus difficile d'éliminer leurs concurrents. Notre système de traitement du langage est sensible à la fréquence des mots individuels mais aussi au nombre de mots qui partagent la même syllabe initiale.

Le rôle de la fréquence se manifeste également dans la lecture et l'orthographe, quand il s'agit de relier les graphèmes et leurs phonèmes en lisant à haute voix. Ellis (2002 : 150-151) soutient que dans la mesure où les lecteurs sont capables de construire une prononciation correcte de nouveaux mots, y compris des « non-mots » (*nonwords*), ils sont capables d'appliquer des « règles » sublexicales de mappage qui relient les graphèmes et les phonèmes.

Cela prend moins de temps aux anglophones adultes de lire des mots ayant des correspondances régulières entre l'écriture et le son et ils le font avec moins d'erreurs que lorsqu'ils lisent des mots ayant des correspondances exceptionnelles. Lors du développement

du langage, les mots à l'orthographe plus rare sont acquis plus tard que ceux à l'orthographe plus régulière, tout comme lors du développement du langage, les mots réguliers qui se conforment aux règles (*pink*) sont acquis plus tôt que les mots ambigus (*cove* vs. *love*, *move*).

Ellis (2002 : 150-151) souligne aussi que parmi les mots irréguliers, ceux qui partagent la même rime et qui sont tous prononcés de la même façon (*look*, *book*, *cook*, etc., ou *calm*, *balm*, *palm*) sont acquis plus tôt que les mots ambigus (*love*). Ces effets de régularité et de cohérence sont plus visibles avec les mots de plus basse fréquence qu'avec les mots de haute fréquence. Les effets de la cohérence et de l'ambiguïté sur la correspondance entre écriture et son ont été simulés avec succès dans les modèles connexionnistes et computationnels.

Enfin, l'effet de la fréquence se manifeste dans la rapidité et la qualité de l'acquisition de l'écriture et dans la lecture. Selon Ellis (2002 : 151-152), l'hypothèse d'une transparence orthographique prédit que les enfants apprenant à lire une langue à orthographe plus transparente que celle de l'anglais ou du français, telle que celle de l'italien ou de l'espagnol, dans laquelle les mappages entre sons et symboles sont réguliers et cohérents, devraient apprendre à lire plus vite que ceux apprenant une orthographe plus opaque où les indices de prononciation sont plus ambigus.

En effet, les enfants espagnols, par exemple, arrivent à lire (en espagnol) une plus grande quantité à partir d'un échantillon de huit mots monosyllabiques et huit mots dissyllabiques (de deux syllabes) que des enfants français ou anglais. Les enfants turcs arrivent à lire et écrire à un niveau de précision plus élevé que les enfants d'autres langues dès la fin de leur première année d'école primaire (qu'on appelle en France le Cours Préparatoire ou CP). Lorsque les statistiques sont plus simples, le temps de traitement est plus court.

En ce qui concerne le lexique, Ellis (2002 : 152-153) explique que la reconnaissance et la production de mots dépendent de leur fréquence d'occurrence dans une langue. Dans les tâches de reconnaissance lexicale de langues écrites, les sujets reconnaissent et nomment plus

rapidement les mots à haute fréquence que les mots à basse fréquence et ils les épellent avec moins d'erreurs. La précision et la rapidité de reconnaissance auditive des mots est supérieure pour les mots à haute fréquence que pour les mots à basse fréquence dans tous les exercices de perception de la parole (lecture, tâche consistant à nommer des objets et perception des signes) et de production lexicale (parler, taper, écrire et utiliser la langue des gestes) chez les enfants et les adultes, en L1 comme en L2. Dans ces modèles, les souvenirs sont formés à partir de toute expérience d'input dans le discours et ces traces retiennent les détails, par exemple, les caractéristiques de la voix du locuteur et le contexte.

Comme l'indique Ellis (2002 : 153), de nombreuses études en morphosyntaxe ont mis en relief le rôle de la fréquence, de la productivité des structures et du modèle de la concurrence dans le cadre du MBU. Une plus grande ambiguïté dans les mappages entre formes et fonctions réduit l'apprentissage à cause d'un plus grand degré de concurrence entre les indices de cette langue.

Ellis (2002 : 153) se réfère entre autres à Taraban et Roark (1996) qui ont manipulé l'ambiguïté dans le mappage des formes nominales avec le genre, dans deux groupes de noms en français. Ces derniers ont montré qu'apprendre le même groupe de noms féminins prenait plus longtemps si les noms du groupe masculin étaient en tant que groupe plus ambigu dans les mappages de genre à partir des terminaisons de ces noms. Cela illustre le fait que non seulement la présence de marqueurs non transparents affecte la rapidité avec laquelle les items non transparents eux-mêmes sont acquis, mais cela ralentit aussi l'apprentissage du système dans tout son ensemble.

Ellis (2002 : 153) fait également référence à l'étude de Rumelhart et McClelland (1986) selon laquelle il est possible de créer un modèle reproduisant l'acquisition du temps du passé des jeunes enfants anglophones. Le modèle a permis de reproduire l'apprentissage en des enfants et montre une tendance à surgénéraliser notamment les verbes irréguliers. Ce



modèle a eu un impact important en construisant les fondations d'une approche connexionniste concernant les recherches en acquisition des langues.

En ce qui concerne l'évolution des langues, Ellis (2002 : 172) se réfère aux études de Bybee (1995), Bybee et Thompson (2000), et Bybee et Hopper (2001) pour montrer que la productivité des modèles phonologiques, morphologiques et syntaxiques dépend surtout de la fréquence des constructions plutôt que de la fréquence des mots. Les mots à haute fréquence permettent « l'enracinement ou l'ancrage » (*entrenchment*) ou la conservation des formes irrégulières et des idiomes ; les formes irrégulières ne survivent que grâce à leur haute fréquence.

Selon Ellis (2002), la fréquence a aussi des implications sur l'enseignement des langues. Pour que les apprenants puissent créer des généralisations correctes et précises, ils doivent avoir entendu et répété de très nombreuses fois des constructions et items :

For language learners to be accurate and fluent in their generalizations they need to have processed sufficient exemplars that their accidental and finite experience is truly representative of the total population of language of the speech community in terms of its overall content, the relative frequencies of that content, and the mappings of form to functional interpretation. The enormity of the lexical pool, the range of frequencies from 60,000 per million down to 1 per million and below, and the wide range of different linguistic constructions, when considered from the point of view of sampling theory, makes it clear that the necessary representative experience for fluency must be vast indeed. (Ellis, 2002 : 167)

En outre, d'après Ellis (2002 : 173), une grande partie de l'acquisition d'une langue consiste en un renforcement progressif d'associations entre des éléments co-occurents. La performance langagière consiste ainsi à exploiter des connaissances probabilistes. Toutefois, pour bénéficier de la fréquence, encore faut-il avoir prêté attention à l'input et plus particulièrement, à certains points spécifiques de l'input. Schmidt (1990) avait observé que pour qu'une structure soit acquise, il fallait d'abord que le locuteur y fasse attention (« *noticing* » *hypothesis*). Schmidt (1993), a par la suite précisé son hypothèse en disant qu'il importe donc non seulement de faire attention à l'input, mais surtout de faire attention à ce qui est pertinent pour un élément particulier de l'acquisition.

Ellis (2002 : 175) est d'accord avec Schmidt (1990, 1993) que si l'apprenant ne se concentre pas sur des formes particulières, il est peu probable que l'apprentissage ait lieu et que l'acquisition soit précise et correcte. Si le locuteur ne fait pas attention à certains points particuliers, l'acquisition va être plus lente, ou risque de ne pas avoir lieu du tout.

Enfin, Ellis (2002 :178) conclut que même si les apprenants analysent l'input de la langue auquel ils sont exposés et y prêtent attention, la fréquence n'est pas la seule explication; d'autres facteurs jouent un rôle majeur : la régularité morpho-phonologique, la complexité sémantique d'une structure, le nombre de concurrents, sa saillance et d'autres facteurs. La fréquence est un facteur important, mais n'est certes pas le seul dans l'acquisition d'une structure.

### **1.3. Autres études liées au Modèle Basé sur l'Usage**

Des études récentes ont mis en relief le rôle de l'input (sa qualité et sa quantité) dans l'acquisition d'une L1 et d'une L2.

#### **1.3.1. Le rôle de l'input : un input riche**

La fréquence de l'input joue un rôle déterminant dans la compréhension, la production et le développement de catégories et règles linguistiques. Diessel (2007) en explique le rôle dans l'acquisition de la langue maternelle des enfants, et précise que l'input de l'enfant est plus riche qu'il n'y paraît, permettant un développement langagier efficace, basé sur un système distributionnel de probabilités. La langue présente dans l'environnement de l'enfant fournit toutes les informations qui lui sont nécessaires afin d'apprendre les parties du discours, et ce, même si on enlevait les informations sémantiques et pragmatiques.

Diessel (2007) indique que non seulement les caractéristiques phonologiques

renforcent les informations que les enfants peuvent extraire des régularités distributionnelles, mais que ces caractéristiques semblent aussi être particulièrement puissantes dans les domaines lexicaux dans lesquels les informations distributionnelles ne sont pas aussi accessibles. Les éléments distributionnels sont utiles pour la catégorisation d'items à haute fréquence et les enfants les rencontrent souvent dans le même contexte. Par contre, ils sont moins utiles pour la catégorisation de mots à plus basse fréquence car ceux-ci ne sont pas assez fréquents pour être associés à un contexte syntaxique particulier.

Toutefois, selon Diessel (2007), étant donné que les items de basse fréquence ont tendance à être plus longs et phonologiquement plus complexes que ceux à haute fréquence, ils fournissent davantage d'informations phonétiques, ce qui peut aider les enfants à créer des catégories linguistiques même pour les expressions les moins fréquentes. Il apparaît donc que la langue entendue dans l'input de l'enfant fournit de riches informations que l'enfant peut utiliser pour apprendre les catégories grammaticales. Par exemple, l'input riche dont bénéficient des bébés de 8 mois leur permet de déterminer les probabilités transitionnelles, telle que la probabilité qu'un élément X apparaisse après Y. Ces informations statistiques jouent un rôle clé dans le développement grammatical. Les enfants développent une grammaire fondée sur leur expérience langagière. Les catégories et les structures linguistiques sont associées à des valeurs probabilistes qui sont déterminées par leur fréquence relative dans l'utilisation de la langue.

Diessel (2007) mentionne plusieurs études qui mettent en relief le rôle de la fréquence, notamment celle de Juliano et Tanenhaus (1993). Ces chercheurs ont analysé le temps de réponse des sujets d'après la place du mot *that*. 89% de toutes les occurrences (*tokens*) de *that* sont des démonstratifs s'ils sont placés en début de phrase, et seulement 11% sont des conjonctions ; après des verbes, 93% des occurrences sont des conjonctions et seulement 7% sont des démonstratifs. Lorsque la fonction et la place de *that* ont été changées et *that* a été

placé dans des contextes moins fréquents (en d'autres termes, lorsque la place du mot ne correspond pas à la fonction la plus fréquente à cette place, et va donc à l'encontre de ce que le locuteur rencontre le plus fréquemment) l'analyse est moins prévisible, ce qui explique pourquoi le temps de réponse des sujets a été plus long que lorsque *that* a été placé dans une position qui correspond à sa fonction la plus fréquente.

Diessel (2007) précise que les éléments sont codifiés dans un système hiérarchique basé sur la fréquence. Les éléments les plus fréquents se trouvent en haut de la hiérarchie et sont accessibles en priorité, ce qui demande aux locuteurs le minimum de temps alors que les éléments moins fréquents se trouvent en bas de la hiérarchie et demandent donc plus de temps à retrouver et à analyser.

MacWhinney (1978, 1982, 1987, 1988) et Bates et MacWhinney (1982, 1987) ont développé ce principe de hiérarchie avec le « Modèle de la Concurrence » (*the Competition Model*). Selon ce modèle, la fréquence joue un rôle primordial dans la compréhension d'une séquence, d'une structure ou d'une phrase. Les utilisateurs d'une langue mémorisent une énorme quantité d'informations statistiques, ce qui détermine leurs attentes linguistiques, et joue un rôle important dans la résolution d'ambiguïtés syntaxiques.

L'acquisition des formes d'une langue résulte de la concurrence entre des items lexicaux, des formes phonologiques et des modèles syntaxiques basés sur l'acquisition correcte (*accurate recording*) de nombreuses occurrences de séquences bien formées entendues dans des contextes, des structures et des modèles différents.

Enfin, selon Inglis (2004), en accord avec les modèles précédemment mentionnés, les séquences qui se trouvent en haut de la hiérarchie, qui ont le plus été entendues et que le locuteur a le plus utilisées deviennent, selon les termes de Langacker (1987), « *entrenched* » c'est-à-dire « ancrées » ou « enracinées » (notre traduction). L'acquisition d'une langue consiste donc à « enraciner » et élargir ces unités, et à en construire de nouvelles.

Selon le rapport qu'Inglis (2004) fait de Barlow et Kemmer (2000), des recherches sur le MBU ont comparé des groupes de différentes langues, des enfants de différents âges, des adultes, des aphasiques, des apprenants d'une L2 et des bilingues. Ces recherches ont montré que les représentations linguistiques sont fortement lexicalisées, l'item lexical étant celui qui contrôle l'analyse grammaticale et l'acquisition en général.

D'après le modèle que décrit Inglis (2004), le traitement et l'acquisition peuvent être bloqués si l'encodage des informations phonologiques est inadéquat ou limité par des bruits (s'il y a quoi que ce soit qui empêche l'input d'arriver correctement à l'oreille de l'apprenant). Par exemple, l'acquisition peut être bloquée ou fortement limitée si l'encodage des bases des constructions sémantiques est inadéquat, si des problèmes dans la rapidité ou la précision dans l'accès au lexique ou des problèmes à mettre en relation l'input et le lexique empêchent le locuteur d'utiliser le lexique pour segmenter la parole entendue, ou si l'apprenant ne peut se concentrer, a des problèmes d'attention ou des problèmes sociaux. Si l'une de ces situations se produit de façon régulière, cela conduira à un mappage incorrect des représentations abstraites et des items initiaux.

Si l'enfant ne reçoit pas le feedback adéquat ou si l'enfant a des problèmes à traiter ce feedback, cela pourrait être insuffisant pour développer un mappage correct et ne permettrait pas d'annuler les mappages incorrects. Des difficultés à mémoriser ou à recevoir l'input pourraient interférer avec la détection des indices et des signaux ; les bruits dans le système peuvent participer au développement d'un grand nombre d'erreurs dans le discours, même si certains mécanismes ont été acquis. Des difficultés à contrôler la sérialisation d'un système parallèle pourrait conduire à la production d'erreurs dans la parole, au bégaiement et à d'autres problèmes dans l'output.

Un mauvais fonctionnement minimal dans l'un de ces domaines pourrait ralentir l'acquisition ; toutefois, étant donné que ces domaines s'entraident, l'enfant ne devrait avoir

qu'un handicap léger qui pourrait être corrigé. Par contre, des blocages majeurs dans l'une de ces habiletés telles qu'un blocage auditif grave pourraient avoir pour conséquence des déformations graves dans le développement du langage.

Cela a, à notre avis, des implications sur l'acquisition d'une langue seconde. Dans une certaine mesure, l'encodage des informations phonologiques en langue seconde/étrangère est partiellement inadéquat et certainement limité à la salle de classe. En outre, l'input n'est pas toujours bien formé si les enseignants, seule source d'input, ne sont pas de langue maternelle et ne maîtrisent pas le système phonologique de la langue cible. Enfin, il n'est pas garanti que l'apprenant entende bien la langue cible, s'il n'a pas reçu au préalable une formation en phonétique. En se concentrant sur le message et en particulier sur le vocabulaire, il risque de ne pas faire attention aux éléments qui ne sont pas nécessairement une entrave à la compréhension en salle de classe, tels que certains aspects phonologiques. Par exemple, si un apprenant prononce la consonne finale non-enchaînée dans *vous avez* et dit /vuz ave/, l'enseignant et tous les élèves de la classe comprendront ce qu'il dit.

Ce n'est donc pas forcément la compréhension (ou le manque de compréhension) du message qui va favoriser l'acquisition d'une structure mais plutôt l'attention sur la forme et sur les « détails » qui ne sont pas toujours indispensables à la compréhension du message. Comme le mentionne Schmidt (1980, 1993), si l'apprenant ne fait pas attention à certains aspects particuliers, l'acquisition ne peut avoir lieu.

Ainsi, pour qu'une acquisition « normale » ait lieu, plusieurs conditions sont nécessaires, ce qui n'est pas toujours le cas en acquisition d'une L2. Le locuteur doit au moins pouvoir bénéficier de la fréquence et de la richesse d'un input « bien formé », avoir une oreille « spécialisée » en phonétique de la L2 qui lui permette d'entendre des différences phonologiques ; il doit être capable de faire attention à certains aspects (phonétiques), et être formé à les produire, grâce notamment aux corrections de locuteurs natifs ou quasi natifs.

### 1.3.2. Fréquence et évolution de la langue

La fréquence d'utilisation des mots et des structures conditionne l'acquisition d'une L1 et d'une L2 mais conditionne aussi l'évolution de la langue : l'évolution des mots, groupes de mots, expressions et règles morphosyntaxiques et phonétiques. La fréquence conditionne l'évolution de tous les aspects linguistiques lexical, syntaxique, morphologique et phonétique.

Comme l'explique Bybee (2001), à un niveau lexical, les mots les plus fréquemment utilisés ensemble ont tendance à « fusionner » et à ne pas être séparés par d'autres mots tels que des adverbes. Ces séquences de mots fréquemment utilisées sont souvent réduites phonétiquement. Les expressions fréquentes sont souvent utilisées dans leur forme réduite, parce que les locuteurs ont l'habitude de les entendre et de les produire. Utiliser des formes réduites que tout le monde a l'habitude d'entendre n'entrave donc pas la communication.

Bybee et Scheibman (1999) ont mis en relief le rôle de la fréquence sur l'évolution de la langue et notamment sur les formes réduites telles que *don't*. Les réductions ou assimilations dépendent uniquement du contexte et d'aucune autre caractéristique phonétique. Les groupes de mots étant souvent utilisés ensemble deviennent des constituants, ce qui facilite les effacements et les réductions du type *don't, I dunno, gonna, wanna*, etc.

De même, selon Bybee et Scheibman (1999 : 577-578), une séquence fréquemment répétée devient automatisée et devient une unité. Les changements phonétiques touchent les mots à haute fréquence plus souvent et plus vite que les mots à basse fréquence. Bybee (2002: 216) précise que les réductions articulatoires surviennent dans les mots et phrases qui sont les plus fréquents, que ce soit avec l'effacement du schwa en anglais, l'effacement du t/d ou l'effacement du d intervocalique en espagnol. Ce sont dans les séquences les plus fréquentes qu'on trouve le plus de réductions et d'assimilations. Lexique, morphologie et phonologie sont liés puisque la prononciation de certains phonèmes dépend de morphèmes et d'autres unités en rapport avec la fréquence.

Diessel (2007) souligne que la fréquence n'est pas seulement une force motrice dans la réduction et la grammaticalisation, elle peut aussi être une force conservatrice et, selon Bybee et Thompson (1997), les expressions fréquemment utilisées résistent souvent aux changements par analogie. Par exemple, en anglais, il y a eu de fortes pressions pour régulariser des formes verbales irrégulières. Près de 200 verbes ont perdu la voyelle de leur racine alternante et ont adopté une forme régulière pour le temps du passé. Si l'on considère les verbes qui sont encore irréguliers aujourd'hui en anglais, on voit que la plupart d'entre eux sont des verbes très fréquents. Leur utilisation fréquente a renforcé leur représentation dans la mémoire, ce qui explique pourquoi ils ont résisté au changement par analogie.

L'effet de conservation de la fréquence d'occurrence des mots ne se restreint pas seulement à des mots isolés mais peut aussi avoir un effet sur les structures syntaxiques. Par exemple, Bybee et Thompson (1997) indiquent qu'on n'utilise pas *do* avec les auxiliaires et les modaux parce que ces verbes ont été entendus et utilisés si fréquemment qu'ils n'ont pas été touchés par le nouvel ordre des mots apparu il y a quelques siècles. En outre, étant donné que les pronoms personnels sont plus fréquents que tout autre type de sujet, les contractions apparaissent généralement avec les pronoms.

Il est d'ailleurs intéressant de remarquer que les liaisons sont obligatoires avec les pronoms clitiques sujets alors qu'elles sont interdites avec d'autres types de sujet, pour les mêmes raisons qu'il y a des contractions avec les pronoms en anglais. Les pronoms sont plus fréquents que les autres sujets, et à cause de cette différence de fréquence, les pronoms sont traités différemment des autres sujets. Nous voyons donc le rôle de la fréquence pour les mêmes éléments (ici les pronoms), en français comme en anglais.

Selon Diessel (2007 : 118-123), la fréquence d'occurrence des items a deux effets apparemment contradictoires: d'une part elle participe à la réduction phonétique et au développement de nouvelles formes ; d'autre part, elle peut être une force conservatrice qui



empêche les structures à haute fréquence d'être nivelées par l'effet d'analogie. Les formes irrégulières les moins fréquentes sont régularisées car elles sont difficiles à mémoriser ; les cas de suppléance (*go-went*) n'apparaissent que dans les formes les plus fréquentes.

De plus, Bybee (2001: 5) remarque que les mots à haute fréquence sont acquis plus rapidement que ceux à basse fréquence et les formations morphologiques irrégulières à haute fréquence ont tendance à garder leurs irrégularités justement à cause de leur haute fréquence. Par exemple, *broke* (le passé de *break*), est apparu 66 fois par million de mots dans l'étude de Francis et Kucera (1982), alors que *damaged* (le passé de *damage*) est apparu 5 fois par million de mots dans le même corpus. Ainsi, l'item *broke* est beaucoup plus fréquent que *damaged*, ce qui suggère que les verbes irréguliers restent irréguliers parce qu'ils sont plus fréquents que les verbes réguliers.

Bybee (2001 : 12) observe que les passés irréguliers ayant des alternances résistent au nivellement analogique. Alors qu'en anglais, les irréguliers *weep/wept*, *creep/crept*, mots à basse fréquence, ont tendance à se régulariser en *weeped* et *creeped*, ceux à haute fréquence, tels que *keep/kept*, *sleep/slept* ne manifestent aucune tendance similaire. Les irrégularités morphologiques ont été préservées par la fréquence et continuent donc à l'être.

La fréquence d'occurrence d'une structure (*Type Frequency*) semble être plus importante que la fréquence des items. Par exemple la flexion du passé en -ed comme dans *damaged* apparaît dans des milliers de verbes et c'est la structure du passé la plus fréquente. Par contre, la structure du passé *broke* en o + cons de l'infinitif + e est beaucoup plus rare que -ed et ne se trouve que dans une classe restreinte *spoke, wrote, rode...* Par conséquent, on peut parler de « productivité » d'une structure. Plus cette structure regroupe d'items, plus elle est forte, plus elle est accessible et prête à être utilisée pour de nouveaux items. La structure de *damaged* est plus fréquente que la structure de *broke*, puisque le passé en -ed regroupe un plus grand nombre d'items que les passés irréguliers. (Bybee, 2001 : 13)

Les structures productives et donc régulières (puisqu'elles regroupent le plus grand nombre d'items, elles deviennent la règle), sont moins touchées par l'effet de fréquence. Ce sont justement les structures moins fréquentes qui sont le plus touchées par l'effet de fréquence des items de cette catégorie. Ainsi, *broke* est un item fréquent qui appartient à une catégorie peu fréquente. La fréquence de l'item compense la fréquence peu élevée de la structure et son manque de productivité. Dans ce cas, la fréquence de l'item joue donc un rôle déterminant.

La fréquence d'utilisation renforce des représentations linguistiques dans la mémoire, ce qui influence à son tour leur activation et leur interprétation dans l'utilisation du langage. La fréquence renforce des attentes linguistiques, puisque les expressions sont arrangées dans un ordre récurrent. Le locuteur s'attend à trouver un mot ou une expression après un certain mot ou une certaine catégorie grammaticale ou sémantique, ce qui influence la compréhension et la production d'unités linguistiques et peut donner naissance à des changements diachroniques.

Enfin, la fréquence favorise le développement de tronçons préfabriqués (*automated chunks*). Les mots se trouvant régulièrement ensemble deviennent des groupes préfabriqués dans lesquels les limites deviennent floues et le groupe entier est réduit. Bybee (2002 : 217-220) précise que les locuteurs natifs stockent en mémoire les mots et leur fréquence d'occurrence et de cooccurrence pour savoir dans quels cas les changements sont les plus appropriés. Si l'on considère les cas d'alternances et de variation à travers les langues, on voit que les locuteurs répètent et reproduisent ce qu'ils ont entendu, les variantes les plus fréquentes dominent les catégories qui ont été formées et sont utilisées dans un plus grand éventail de contextes, alors que les variantes moins fréquentes ont un statut marginalisé et finissent par disparaître.

Comme le conclut Diessel (2007 : 123-124), bien que la fréquence des items et des structures joue un rôle majeur dans l'acquisition, l'utilisation et l'évolution des langues, les phénomènes d'analogie et d'autres phénomènes similaires en rapport avec la fréquence contribuent à la continuation de schémas linguistiques, à la formation d'expressions nouvelles et au nivellement des formes irrégulières aboutissant à des régularisations. Les connaissances linguistiques sont influencées par des pressions communicatives et cognitives.

### 1.3.3. Fréquence, connexionnisme et variationnisme

La langue évolue en grande partie à cause de l'usage que les locuteurs en font, et l'acquisition d'une L1 et d'une L2 est fortement influencée par la fréquence de certains mots et de certaines structures. Toutefois, d'autres facteurs sont également importants dans l'acquisition d'une L1 ou d'une L2 : la saillance<sup>9</sup>, la pertinence et la récence<sup>10</sup> qui sont toutes en rapport avec le contexte d'acquisition et de communication. Un élément *saillant* est un élément que le locuteur peut remarquer facilement, un élément *pertinent* est ce qui est utile pour le locuteur et un élément *récent* est ce que le locuteur a entendu et/ou utilisé récemment.

Tarone (2002 : 287) suggère que la langue utilisée par les locuteurs varie, et ces variations dépendent de plusieurs facteurs dont la fréquence. Puisqu'on connaît ces facteurs, il est possible de considérer l'acquisition d'une L1 ou d'une L2 dans un cadre variationniste (voir par exemple, Tarone, 1985; Bayley et Preston, 1996), en accord avec les sociolinguistes variationnistes et les chercheurs en acquisition des langues secondes (Bayley, 1994; Broner, 2001; Parrish et Tarone, 1988; Tajika, 1999; Tarone, 1979, 1985, 1988, 1998; Young, 1991) et en accord avec le modèle psycholinguistique de variation linguistique proposé par Preston (2000).

---

<sup>9</sup> La saillance (saliency en anglais) : 1. Qualité d'être saillant. 2. Grande probabilité d'être remarqué. D'après <http://fr.wiktionary.org/wiki/saillance>

<sup>10</sup> La récence (recency en anglais) : caractère de ce qui est récent. D'après <http://www.larousse.fr/dictionnaires/francais/r%C3%A9cence/66947>

Il s'agit alors de déterminer les facteurs dont dépendent les variations pour lesquelles il y a le choix entre deux formes ou de trouver les implications de l'utilisation d'un ou de plusieurs facteurs, le tout ayant des influences probabilistes. Les variantes qui ne sont utilisées que dans des conditions particulières (en langue recherchée) sont plus difficiles à utiliser correctement parce qu'elles sont moins « enracinées ». Comme le mentionne Tarone (2002 : 291) : « We are more fluent in our vernacular. Other language forms (native or nonnative) are postvernacular, less deeply embedded, and weaker. We are always less fluent in the postvernacular. The notion that language constructions internalized in the postvernacular period are always weaker parts of the grammar relates to the traditional notions of style and attention in sociolinguistics. » Selon Tarone (2002 : 291-292), certains facteurs tels que l'attention et la pertinence d'un mot, d'une séquence ou d'une structure sont aussi importants.

L'interaction en contexte de communication authentique va « ancrer, enraciner » les structures en cours d'acquisition. Les intentions et les conditions de l'apprenant rentrent en compte ; la fréquence, la récence et le contexte activent un schéma. Si l'apprenant fait attention, s'il est sélectif et remarque certaines structures plus que d'autres, si elles sont plus importantes pour lui, même si elles sont moins fréquentes dans l'input, alors l'apprenant a plus de chances de maîtriser ces structures. Lors d'interactions, l'apprenant obtient non seulement l'input linguistique dont il a besoin mais aussi l'occasion d'activer des schémas et de les renforcer.

Enfin, Tarone (2002 : 293-294) souligne que lorsque les enfants jouent avec la langue, ils travaillent souvent les variantes de la langue, en particulier les moins fréquentes dans des jeux d'imitation et de parodie. La langue des enfants est en constant changement car les enfants sont plus flexibles et plus sensibles aux variations phonétiques ainsi qu'aux autres variations linguistiques.

L'étude de Hoff (2002) met en évidence le rôle de l'input dans l'acquisition et la production du vocabulaire d'enfants de 2;5 ans à 4;5 ans. Même si toutes les mères de cette étude appartiennent à la classe moyenne, elles appartiennent à différentes couches de la classe moyenne. L'étude suggère que la quantité et la qualité de l'input des mères diffèrent selon leur statut socioéconomique dans la langue qu'elles utilisent en s'adressant à leur enfant. La parole de la mère a ensuite des répercussions sur l'acquisition et la production du vocabulaire des enfants. La production des enfants est donc directement liée à l'input reçu.

Dans cette même étude, les mères de milieu socio-économique plus élevé parlent plus souvent et plus longtemps à leurs enfants et élicitent plus de réponses de leur part, notamment en leur posant plus de questions. Elles créent des situations langagières plus variées et passent plus de temps à lire à leur enfant ce qui entraîne des activités en rapport avec le vocabulaire telles que nommer des objets, reformuler une histoire, poser des questions et y répondre. Cela a aussi pour conséquence que l'enfant se familiarise avec une langue et une prononciation plus recherchées. De plus, la lecture régulière de contes permet des répétitions de mots et de structures. Par exemple, de nombreuses histoires commencent par « il était une fois, une petite fille qui / un petit garçon qui ». La lecture de livres différents, permet également une exposition à un vocabulaire plus varié.

L'étude de Hoff (2002), met en relief le rôle de l'entourage dans la qualité et la rapidité d'acquisition du vocabulaire. Les enfants issus d'un milieu socioéconomique plus élevé peuvent nommer plus de mots que les enfants de milieu socioéconomique plus bas du même âge. Les différences de production des enfants de différents milieux se manifestent à tous les âges, et jusqu'à l'âge le plus avancé de l'étude, c'est-à-dire jusqu'à l'âge de 4 ans.

Nous voyons ici certains aspects des théories connexionnistes selon lesquelles les expériences langagières des enfants ont des répercussions directes sur leur acquisition.

#### **1.3.4. Recherches en reconnaissance de la parole**

Le rôle de l'input et de la fréquence est aussi utilisé dans les programmes informatiques d'intelligence artificielle et de reconnaissance de la parole qui se basent sur les résultats des recherches de spécialistes sur ce qui se passe dans le cerveau humain lors du traitement des informations. Si l'on comprend ce qui se passe dans notre cerveau, on peut mieux le reproduire de manière artificielle. Si un ordinateur peut reproduire certains comportements d'après les données fournies, cela suggère que les théories sont validées. On peut donc utiliser ces résultats à d'autres fins, notamment pour l'enseignement des L2.

Gass et Mackey (2002 : 252-253) expliquent que les systèmes initiaux dans le domaine de la reconnaissance de la parole étaient symboliques ou gouvernés par des règles et ont été dérivés à partir des intuitions des créateurs du système au sujet de la langue. Mais depuis les trente dernières années, le domaine de la reconnaissance de la parole s'est concentré sur des systèmes basés sur la fréquence, sur des modèles statistiques de données empiriques, notamment des modèles de fréquence de séquences de trois mots, appelées trigrammes (Jelinek, 1976). Ces systèmes se sont avérés beaucoup plus précis et adaptables que ceux qui étaient basés sur des règles (notamment syntaxiques).

Gass et Mackey (2002) en concluent que la tendance récente en reconnaissance de la parole consiste à développer les méthodes basées sur la fréquence avec des structures plus riches, et des structures capables de généraliser et d'établir des liens avec des éléments linguistiques plus éloignés. La performance de ces modèles peut être améliorée en employant un système basé sur la fréquence à plusieurs niveaux et en utilisant des structures plus riches que de simples séquences de mots de surface. Selon ces derniers, la performance du système dépend grandement de la quantité de l'input, mais en se concentrant sur des structures plus saillantes, l'apprentissage basé sur la fréquence est devenu plus précis et plus efficace. Le lien

entre la saillance, la fréquence et les perceptions des modèles de l'input est également important pour l'apprentissage d'une L2.

Dufeu et al. (2006) ont établi un réseau de neurones artificiels pour simuler l'apprentissage des formes orthographiques des mots du français chez les apprentis lecteurs. Les chercheurs ont d'abord élaboré la fréquence des mots que rencontrent les enfants de 6 à 11 ans à partir de leurs manuels scolaires. De cette façon, ils ont déterminé un effet de fréquence lexicale (les mots de haute fréquence sont reconnus plus aisément que ceux de basse fréquence) ainsi qu'un effet de voisinage orthographique. La différence entre la reconnaissance de mots fréquents et moins fréquents est très forte au début (82% en 1<sup>e</sup> année contre 50.9% en 4<sup>e</sup> année), alors que la fréquence dans le voisinage orthographique joue un rôle moins important.

Selon Dufeu et al (2006), la tâche de décision lexicale consistant à manipuler la fréquence des mots et le voisinage orthographique montre que les mots qui apparaissent le plus fréquemment dans une langue sont perçus plus rapidement et produisent moins d'erreurs de reconnaissance que des mots qui apparaissent moins fréquemment.

Entraîné à partir des mots vus par les enfants, ce modèle (que les auteurs nomment « carte auto organisatrice ») a généré des résultats similaires à ceux des enfants. Ce modèle peut reproduire les effets de fréquence lors de l'acquisition de la lecture. Les performances du modèle se rapprochant de la performance des enfants, mettant en valeur le rôle important des facteurs de fréquence et de voisinage. Les résultats renforcent l'hypothèse selon laquelle la construction du lexique orthographique chez les enfants est un processus guidé principalement par les données auxquelles ils sont exposés.

Si l'on peut enseigner à un ordinateur comment retrouver et utiliser des informations, de façon similaire à des enfants L1 de 6 à 11 ans, peut-on enseigner à des apprenants L2 le vocabulaire, et la langue en général, telle que l'apprennent les enfants L1?

Les études ci-dessus mentionnées suggèrent qu'un ordinateur peut reproduire, dans une certaine mesure, l'apprentissage des enfants L1, grâce à la fréquence et au MBU. Des résultats semblables ont été obtenus quelles que soient les langues étudiées, et ce dans tous les domaines linguistiques, y compris l'acquisition des langues secondes/étrangères. Il s'avère donc, que si l'on est capable de reproduire avec un logiciel le processus d'acquisition des enfants et les productions enfantines d'après l'input reçu, c'est qu'on comprend ce processus d'acquisition. Puisque les facteurs majeurs sont la fréquence et l'analogie, ces facteurs devraient être tout aussi utiles dans l'acquisition d'une L2.

### **1.3.5. Le Modèle Basé sur l'Usage et l'acquisition d'une L2**

Si le MBU a fait ses preuves dans les études en L1, qu'en est-il des études en L2? De nombreux chercheurs se sont intéressés au rôle de l'input, tels que Wagner-Gough et Hatch (1975), Larsen-Freeman (1976), Hamayan et Tucker (1980) et Lightbown (1983). Gallaway et Richards (1994) ont réuni de nombreuses recherches en psychologie et en psycholinguistique dans leur livre *Input and Interaction in Language Acquisition*. Ces recherches indiquent, une fois de plus, le rôle majeur de la fréquence de l'input dans l'apprentissage d'une langue.

Selon Bardovi-Harlig (1987) la fréquence de l'input peut l'emporter sur les difficultés à apprendre des traits marqués. Gass et Lakshmanan (1991) ont trouvé une corrélation entre les données de l'input et de l'output. Dans son Hypothèse de l'Interaction, Long (1996) a suggéré que l'interaction lors d'une conversation promeut le développement d'une L2, ce que Vygotsky (1978) avait déjà observé chez les enfants L1 avec « *the Zone of Proximal Development* ».

Eskildsen (2008) explique par exemple le développement de l'acquisition de *can* chez un apprenant de l'anglais L2. Certaines formes sont transitoires et réduites à certains



contextes, surtout en début d'acquisition (*can you, can I*) alors que certaines séquences apparaissent et disparaissent en cours d'apprentissage. L'apprenant utilise les formes dont il a besoin et ces formes changent au fur et à mesure de son expérience avec la langue, ce qui rappelle les modèles connexionnistes. Il semble que pour que des structures et des règles générales se construisent, il faut non seulement que le sujet les ait entendues, mais aussi qu'il les ait pratiquées en situation de communication pour que celles-ci « se fixent » et deviennent plus généralisées.

Sergeeva et Chevrot (2008) ont analysé l'influence de la fréquence sur l'acquisition des formes verbales régulières et irrégulières du français chez des russophones adultes. Les formes des verbes à fréquence cumulée élevée sont significativement mieux formées que celles des verbes à basse fréquence cumulée. Le taux de formation juste des verbes plus fréquents est significativement plus élevé que celui des verbes moins fréquents pour les étudiants avancés.

L'effet de fréquence se fait ressentir à un niveau d'apprentissage plus élevé, ce qui indique que pour pouvoir bénéficier de l'effet de fréquence, il faut d'abord avoir entendu et pratiqué certaines structures un grand nombre de fois. Les verbes irréguliers ici testés (*faire, prendre, devoir, boire, craindre, traduire*), appartiennent à des classes restreintes moins fréquentes que des verbes plus réguliers (*aimer, finir*), et il y a une concurrence entre les formes (différentes formes pour les verbes du 3<sup>e</sup> groupe) ce qui ralentit le processus d'acquisition. C'est la raison pour laquelle l'effet de fréquence se fait ressentir à un niveau plus élevé pour les groupes de verbes à conjugaison plus complexe. De plus, dans l'étude, il s'agissait de plusieurs modes et temps (présent, impératif, imparfait, subjonctif présent, passé simple...). Là aussi, on voit le rôle majeur de la fréquence et de l'effet d'analogie.

Lors de son étude de l'acquisition de certaines variantes sociolinguistiques du français par des apprenants anglophones, Rehner (2010a et 2010b) explique en se référant aux

recherches de Mougeon et al. (2010 : 3) que plusieurs phénomènes contribuent à l'acquisition d'une structure : « Mougeon and his colleagues found that the frequency and patterns of use of nearly a dozen variants in the speech of the high school immersion learners depended primarily on the availability of an English equivalent, on the structural complexity of the variants, and on the frequency of the variants in the educational input. »

Rehner (2010b : 308) observe que dans l'étude de Rehner et Mougeon (1999), les élèves d'écoles d'immersion omettent le *ne* de négation à l'oral dans 28% des cas, ce qui se rapproche des productions des enseignants d'écoles d'immersion qui l'omettent dans 29% des cas dans l'étude de Mougeon, Nadasdi et Rehner (2010). Ceci souligne le rôle de l'input dans la production des apprenants L2 qui reproduisent ce qu'ils entendent, et leurs productions se rapprochent de celles de leurs enseignants, leurs principaux modèles linguistiques.

Rehner (2010b : 303-308) mentionne aussi que les omissions de *ne* des apprenants sont plus élevées à l'université (42%) qu'à l'école secondaire après que ceux-ci aient été en contact avec plus de francophones et de cours en français, mais que la production de certaines variantes dépend du contact que les étudiants ont eu avec les francophones. Rehner précise aussi que la langue maternelle joue également un rôle, tout comme le type d'enseignement reçu. Étant donné qu'en espagnol et en italien par exemple, il est impossible d'omettre le *ne*, les locuteurs de ces langues ont une forte tendance à le conserver en français. Enfin, les étudiants ayant été en immersion en omettent plus que ceux qui ont été en français de base, même s'ils ont tous suivis des cours en français à l'université et qu'ils ont tous eu accès à la langue française de façon similaire. Ces cours universitaires ne permettent pas de contrebalancer l'input reçu à l'école.

Les deux études de Rehner (2010a, 2010b) soulignent le rôle majeur de plusieurs facteurs dans l'acquisition de structures en L2 : l'input reçu (et la fréquence de certaines structures dans cet input), la langue maternelle et la complexité d'une structure. Même si ces

deux études analysent des éléments différents (ici l'omission du *ne* et la production d'autres variantes sociolinguistiques), elles s'accordent avec d'autres études telles celle de Gass et Mackey (2002) qui ont analysé le rôle de plusieurs facteurs dans l'acquisition d'une L2 en rapport avec la fréquence de l'input. En L2, la fréquence est nécessaire mais non suffisante à cause du transfert de la L1 qui peut ralentir l'acquisition. Lorsqu'une structure existe en L1, l'acquisition se fait plus rapidement et de manière plus cohérente et régulière que lorsque la structure n'existe pas en L1.

Gass et Mackey (2002 : 253) précisent en se référant à l'étude de Pienemann et Johnston (1986) que, bien que la fréquence joue un rôle majeur en acquisition d'une L2, il y a tout de même des cas où elle ne peut pas expliquer pourquoi certaines structures sont acquises avant d'autres. Par exemple le *s* de la 3<sup>e</sup> personne est fréquent, il est simple à comprendre et pourtant il n'est acquis qu'à la 5<sup>e</sup> étape (sur 6).

À ceci, nous répondons en utilisant le MBU que plusieurs facteurs contribuent à l'acquisition d'une structure : sa complexité, sa fréquence, sa saillance et sa productivité. La productivité d'une structure en détermine la fréquence, par opposition à la fréquence des structures concurrentes, et aussi le nombre d'items qu'elle englobe, par rapport au nombre d'items qu'englobent les structures concurrentes.

Or, si l'on considère le cas du *s* de la 3<sup>e</sup> personne du singulier en anglais, cette structure paraît simple au premier abord (il suffit de rajouter un *s*) et elle est relativement fréquente, puisque présente dans tous les verbes sauf les modaux. Toutefois, c'est une structure peu productive. D'une part, aucune autre personne de la conjugaison du présent n'a de *s* (*I go, you go, etc.*), d'autre part, aucun autre temps ne permet une conjugaison en *s* à la 3<sup>e</sup> personne du singulier (*he will go* et non *he wills go\** ou *he will goes\**). Lorsque nous disons que le *s* du présent est une structure « peu productive », cela signifie, peu productive par rapport aux concurrents. Si l'on considère toutes les occurrences de verbes (toutes les formes

verbales, tous les modes, tous les temps et toutes les personnes) les formes verbales sans *s* sont beaucoup plus fréquentes que les formes avec *s*. Puisque le *s* ne s'utilise qu'à une seule personne d'un seul temps, alors que pour toutes les autres personnes et tous les autres temps, il n'y a pas de *s*, les cas sont beaucoup plus fréquents sans *s* qu'avec *s*. Les concurrents sont largement plus nombreux. C'est donc une structure peu productive.

De plus, le *s* est généralement utilisé pour marquer le pluriel dans tous les autres cas où on utilise un *s* en anglais (*schools, girls*). Autrement dit, au niveau cognitif, la règle n'est pas aussi simple qu'il n'y paraît parce qu'il n'y a aucun autre cas de *s* au singulier, et justement les *s* sont utilisés uniquement pour le pluriel. Si le *s* était utilisé pour *they*, par exemple, elle serait un peu moins complexe pour les apprenants de l'anglais L2/étrangère, parce qu'ils associeraient le *s* au pluriel, tout comme pour le pluriel des noms. Il apparaît donc que la règle est peu productive et relativement complexe ce qui explique qu'elle soit acquise plus tard et que les cas les plus fréquents (par rapport aux concurrents) et les plus réguliers qui n'ont pas de *s* soient acquis plus tôt, par analogie.

De même, Trahey et White (1993) ont remarqué que bien que les apprenants d'une L2 soient capables de reconnaître une forme correcte de la L2, ils ne sont pas toujours capables de remarquer si une structure est incorrecte. Les apprenants remarquent par exemple qu'une structure de la langue cible est différente de leur L1 (place des adverbes en anglais L2, français L1). Ils reconnaissent que c'est une structure grammaticale en anglais. Par contre, ils ne reconnaissent pas une structure fautive si celle-ci utilise les paramètres de leur L1. Cela indique que la fréquence permet de reconnaître la bonne formation de certaines structures si celles-ci sont différentes de la L1, mais ne suffit pas pour éviter les erreurs de transfert de la L1 si les structures sont acceptables en L1 mais pas en L2.

Il faut tout de même préciser que les formes en question n'ont peut-être pas été entendues suffisamment de fois pour que la fréquence compense les différences entre la L1 et

la L2. Si elles ont été entendues mais que les apprenants n'ont pas fait attention à la forme, le rôle de la fréquence peut être minimisé. En fait, soit les structures différentes en L1 et L2 ne sont pas acquises et il s'opère rapidement un phénomène de fossilisation avec transfert de la L1, soit elles sont acquises à plus long terme mais nécessitent une plus grande fréquence que les structures similaires ou les mêmes en L1.

En effet, Spada et Lightbown (1993) ont montré qu'indépendamment de l'input, les formes interrogatives de plus haut niveau ne peuvent être acquises que lorsque les apprenants sont prêts, quand ils atteignent le niveau requis et ont « de la place pour traiter » ces formes. Nous pensons qu'il s'agit ici du même phénomène que pour l'acquisition du *s* de la 3<sup>e</sup> personne du singulier. Lorsqu'on parle de fréquence, il ne s'agit pas seulement de la fréquence d'une structure, mais plutôt du pourcentage de cette structure dans la totalité de l'input, c'est-à-dire que la fréquence d'une structure est liée à sa complexité et à sa productivité. Ce n'est pas parce qu'un locuteur a souvent entendu certaines séquences qu'il va forcément les apprendre avant des séquences moins fréquentes. Si l'on considère l'ensemble de l'input d'un locuteur adulte, nous supposons que les cas de phrases affirmatives sont plus fréquents que les cas de phrases interrogatives et négatives. Parmi ces phrases affirmatives et négatives, certaines sont plus complexes et plus fréquentes que d'autres.

Par exemple si on utilise la base de données de *Lexique*<sup>11</sup> élaborée par l'équipe du CNRS<sup>12</sup> à partir de sous-titres de films réunissant 16.6 millions de mots d'après une moyenne de 4 films (nous décrivons cette banque de données plus en détail dans le chapitre 3), on trouve que *do you* a une fréquence de cooccurrence de 181, *did you* a une fréquence de cooccurrence de 77 et *does he* a une fréquence de cooccurrence de 6 (par million de mots). On voit donc que *do you* est beaucoup plus fréquent que *did you* qui est lui-même beaucoup plus fréquent que *does he*.

---

<sup>11</sup> <http://www.lexique.org/moteur/RechCoOc.php>

<sup>12</sup> Centre National de la Recherche Scientifique : National Center for Scientific Research, France

Pour comprendre l'ordre d'acquisition, il faut comparer des séquences à d'autres séquences, et non considérer uniquement le taux de fréquence d'une structure isolée. On comprend donc pourquoi *do you* sera acquis avant *did you*, qui sera sans doute acquis avant *does he* (qui est non seulement le moins fréquent, mais aussi le moins productif). Cela pourrait être la raison pour laquelle les enfants anglophones acquièrent le *s* de la 3<sup>e</sup> personne beaucoup plus tard que d'autres morphèmes (Lightbown et Spada, 2006 : 3) et apprennent *did* après *do* (Lightbown et Spada, 2006 : 7).

Gass et Mackey (2002 : 254) en concluent que l'existence de séquences développementales n'élimine pas forcément le rôle de la fréquence. Ceux-ci supposent que lorsque les apprenants n'en sont pas encore au niveau développemental suffisant pour utiliser l'input, ils pourraient stocker l'input et l'utiliser plus tard, quand ils sont prêts.

Nous pouvons déduire de ces études que les apprenants de niveau moins avancé n'ont pas encore eu le temps de créer leur inventaire, de créer des classes et des schémas abstraits, et qu'ils apprennent d'abord les structures qui sont les plus fréquentes en rapport avec l'input reçu. Par exemple dans le cas des verbes irréguliers de l'étude de Sergeeva et Chevrot (2008), les apprenants utilisent les structures les plus fréquentes de leur inventaire et l'analogie.

Il en est de même dans l'étude de Rehner (2010b) : l'effet de fréquence se fait ressentir pour les éléments les plus fréquents, distinguant les élèves d'immersion qui ont pu bénéficier de l'effet de fréquence dans leur input des élèves du français cadre, dans leur omission du *ne*, comme dans leur production du *on*, ou d'autres variantes fréquemment entendues. Mais lorsqu'une variante est plus rare, comme par exemple, le verbe *demeurer* (par opposition à *habiter*, *vivre* et *rester*), les apprenants du français de base comme les apprenants de français d'immersion ne l'utilisent pas (Rehner, 2010a :11) et l'effet de fréquence est peu utile puisque ces éléments sont peu fréquents dans l'input reçu pour les deux groupes.

Les résultats des études sur l'acquisition d'une L2 vont dans la même direction que les études d'une L1 chez les enfants. On peut observer de nombreuses similarités entre l'acquisition d'une L1 et d'une L2. Selon le modèle de la hiérarchie et de la productivité des structures, la théorie de la concurrence de Bates et MacWhinney (1987) et MacWhinney (1987), les structures regroupant le plus grand nombre d'items sont acquises avant les structures regroupant moins d'items, les items les plus fréquents sont appris avant les moins fréquents, et les structures les plus productives sont acquises avant les moins productives.

Bien qu'il y ait de nombreuses similarités entre l'acquisition d'une L1 et d'une L2, notamment les étapes par lesquelles passent les apprenants et le rôle de la fréquence et de la productivité de structures dans l'acquisition d'une langue, l'acquisition d'une L1 diffère en quelques points de l'acquisition d'une L2.

Premièrement, comme le fait remarquer Ellis (2002), les enfants acquièrent une L1 en même temps qu'ils développent leurs connaissances sur le monde, alors que les adultes apprenant une L2 construisent leurs connaissances sur des concepts déjà établis. De plus, les adultes ont des moyens sophistiqués de traiter les informations et peuvent apprendre une langue de façon explicite, lors d'exercices conscients de résolution de problèmes et avec des déductions, et ce, beaucoup mieux que les enfants.

Deuxièmement, l'input des adultes est plus restreint que celui des enfants. Les enfants sont exposés de façon naturelle à des situations où les parents, et autres personnes s'occupant d'eux, participent à leur développement en échafaudage (Tomasello et Brooks, 1999), alors que l'enseignement des adultes en classe peut déformer, fausser les modèles d'exposition, les fonctions et les formes d'interactions sociales (Ellis et Laporte, 1997).

Troisièmement, il y a le transfert d'une L1 vers la langue cible. Les adultes apprennent une L2 en basant leurs nouvelles connaissances sur leurs connaissances préexistantes, notamment celles de leur langue maternelle. (Kellerman, 1995; MacWhinney, 1992; Odlin,

1989). Si une structure n'existe pas, ou que celle-ci est différente dans la langue cible, cela peut ralentir ou compromettre le développement de la langue cible (Trahey et White, 1993).

Pour comprendre les phénomènes en jeu en acquisition d'une L2, il faut comprendre les phénomènes communs en L1 et en L2, mais aussi ceux qui sont spécifiques à l'acquisition d'une L2, comme nous venons de le faire remarquer. D'après le MBU, la fréquence et la productivité des structures joue un rôle majeur dans l'acquisition d'une L1 et d'une L2. Les structures et les items les plus fréquents sont ceux qui sont appris en premier, ceux pour lesquels il y a moins d'erreurs, ceux qui sont en haut de la hiérarchie, et ceux qui sont accessibles le plus rapidement. Ces items et structures se composent souvent de séquences préfabriquées. Plus les séquences sont entendues, plus elles sont enracinées. Plus elles sont enracinées, plus elles sont utilisées, et plus elles sont utilisées, plus elles sont enracinées.

Les items les plus fréquents sont reconnus plus rapidement dans des tâches de lecture ou de reconnaissance lexicale, et ceux qui créent le moins de problèmes d'orthographe aux enfants francophones. Les mots et séquences les plus fréquemment entendus sont ceux qui soit résistent aux changements, ce qui est le cas avec les exceptions, notamment les verbes irréguliers ; soit au contraire, ceux qui subissent des transformations, troncations, assimilations ou « fusions », comme avec *don't*, *gonna*, etc. (cf. Bybee et Scheibman, 1999).

Les résultats des études sur l'acquisition d'une L1 dans le cadre du MBU ont été reproduits avec succès dans des logiciels de reconnaissance de la parole et d'intelligence artificielle (Rumelhart et McClelland, 1986 ; Gass et Mackey, 2002 ; Dufeu et al., 2006). Les études en psychologie et psycholinguistique portant sur l'acquisition d'une L1 (Tomasello, 2000 ; Ellis, 2002 ; Inglis, 2004 ; Diessel, 2007) ont donné des résultats similaires en acquisition d'une L2 (Eskildsen, 2008 ; Sergeeva et Chevrot, 2008). Il semble que le processus d'acquisition soit globalement le même, si les conditions d'apprentissage sont



similaires (dans des situations de communication authentique) : la fréquence des mots et des structures en influence leur acquisition.

Il existe toutefois une différence primordiale, c'est que le processus d'acquisition d'une L2 chez l'adulte est aussi influencé par d'autres facteurs tels que le transfert de la L1 sur la L2. Il est possible que cela soit le cas pour différentes raisons : les conditions d'apprentissage des adultes sont souvent différentes et restreintes à la salle de classe ; l'input est plus restreint et peut-être insuffisant pour compenser l'interférence de la L1. Les conditions d'apprentissage des adultes ne sont pas aussi favorables que les conditions dans lesquelles les enfants apprennent une langue (Tomasello et Brooks, 1999).

La L1 joue un rôle à plusieurs niveaux. Si certaines structures sont similaires, elles peuvent être apprises plus facilement et plus rapidement ; d'autre part la L1 peut influencer la rapidité et le chemin que prend l'acquisition (cf. Zobl, 1982). Si certaines structures sont trop différentes de la L1 et sont, de surcroît, complexes, cela peut inhiber ou entraver l'acquisition. Ainsi, même si les apprenants ont souvent entendu des structures, ils les reconnaissent comme correctes, mais ne sont pas forcément capables de reconnaître une structure fautive dans laquelle la structure cible a été faussée par l'interférence de la L1 (Trahey et White, 1993).

Dans un sens, on peut expliquer ce phénomène de la façon suivante. D'après la théorie de la hiérarchie et de la concurrence de MacWhinney (1978, 1982, 1987), et Bates et MacWhinney (1979, 1982), les structures et items les plus fréquents se trouvent en haut de la hiérarchie et sont plus facilement accessibles. Ce qui place ces items ou structures en haut de la hiérarchie, c'est le nombre de fois que les locuteurs les ont entendus et utilisés. Or, si un item n'a été entendu par exemple qu'une vingtaine de fois en classe, alors que l'item correspondant en L1 a été entendu plusieurs centaines ou milliers de fois en L1, il semble logique que l'item de la L1 se trouve en haut de la hiérarchie et soit beaucoup plus accessible que celui de la L2.

Si un enfant apprend une L2, il n'a pas encore eu le temps d'entendre les structures de la L1 des milliers de fois, c'est pourquoi pour un enfant, la concurrence de la L1 est moins forte. De plus, l'enfant est moins « sélectif » que les adultes. Les adultes cherchent à comprendre, en comparant avec leurs connaissances implicites de leur L1. Puisqu'ils savent lire, ils veulent savoir comment s'écrivent les mots, parce que l'écrit les aide à mémoriser.

Par contre, pour les apprenants adultes d'une L2, la prononciation requiert un entraînement à l'écoute et à la production orale de certains phonèmes et éléments suprasegmentaux, car leur oreille s'est spécialisée dans leur langue maternelle et n'est pas accoutumée à la phonétique de la langue cible. Comme l'explique Lauret (2007) :

Au contraire du vocabulaire et de la grammaire, la prononciation (la « forme » de l'oral), est une compétence physique qui recouvre à la fois de la perception (qu'on met facilement en rapport avec l'oreille et parfois même avec l'oreille musicale) : « Ce sont mes oreilles, je n'entends pas la différence entre ces deux sons. » et la production : « C'est un mot impossible à prononcer pour moi. » (Lauret, 2007 : 15)

Des troubles de « perception » causent des lacunes dans l'apprentissage (Barlow et Kemmer, 2000). On peut, dans cette perspective, trouver des similarités entre un apprenant d'une L2 et un apprenant ayant des problèmes d'audition dans une L1. C'est la raison pour laquelle, en L2, d'autres facteurs sont encore plus importants qu'en acquisition d'une L1 pour contrebalancer cela: l'intérêt, l'attention (Schmidt, 1990 et 1993), la saillance, la pertinence, la récence (Tarone, 2002), l'instruction formelle, le feedback et la correction.

Si une structure est peu fréquente et peu « productive », et qu'en plus, l'item à l'intérieur de cette structure est lui-même peu fréquent, et que cette structure est différente ou absente de la L1, le locuteur adulte aura de plus grandes difficultés à maîtriser cette structure.

En effet, si ni la structure, ni l'item ne sont fréquents, alors l'un ne peut pas compenser l'autre. Cela va être d'autant plus le cas si le locuteur a un stimulus peu varié et peu fréquent (par exemple, les francophones minoritaires ne parlant le français qu'à la maison, ou des apprenants n'ayant parlé le français qu'à l'école).

En outre, si le locuteur ne connaît pas de règles formelles l'incitant à contrebalancer la complexité d'une structure, et n'a pas reçu de feedback correctif quant à sa prononciation, alors le locuteur a peu de chances de maîtriser cette structure. Les items « correctement produits » n'ont pas été entendus suffisamment de fois pour que la structure soit « validée », d'où une lacune dans l'apprentissage.

L'acquisition pourrait être incomplète et se restreindre à certains items, tout comme en début d'apprentissage où l'enfant répète certains mots et séquences, sans être pour autant capable d'en généraliser les règles à d'autres situations. Il va y avoir ici de nombreuses variations, étant donné que de multiples facteurs déterminent l'acquisition de l'apprenant.

Enfin, nous pouvons également émettre l'hypothèse que tous ces facteurs participant à l'acquisition sont directement liés. Si un apprenant a entendu beaucoup de français, en milieu majoritaire, il y a plus de chances que des structures aient été plus fréquemment entendues et utilisées et que des items aient été plus fréquents et plus saillants pour le locuteur. L'expérience langagière de chacun, d'après les théories connexionnistes et variationnistes, joue un rôle majeur.

En milieu majoritaire, la langue a plus d'importance, et elle est étudiée de manière formelle et plus systématique. Dans des pays et régions francophones majoritaires, les élèves suivent généralement des cours de grammaire, d'orthographe, de littérature, etc. Tandis que dans des situations où le français est un outil de communication, notamment à l'école, puisque le français n'est ni la langue maternelle des apprenants, ni celle des enseignants, les conditions d'apprentissage sont différentes.

Les apprenants doivent d'abord apprendre à parler la langue, c'est-à-dire apprendre le vocabulaire et la grammaire de base avant d'étudier la littérature et d'autres sujets connexes ; le contenu des cours est, au moins au début, différent des cours de français de francophones majoritaires. De plus, l'input reçu est restreint à la salle de classe, puisque les locuteurs

parlent leur langue maternelle à la maison, et en dehors de l'école ou l'université. Enfin, les locuteurs subissent l'effet de leur langue maternelle dans leur production de la L2.

Cela veut dire que les francophones majoritaires bénéficient de conditions d'apprentissage nettement avantageuses à tous les niveaux linguistiques, ayant un input plus riche et plus varié. C'est pourquoi la fréquence d'une structure ou d'un item ne devrait pas avoir le même impact sur différents groupes linguistiques, vu que leurs conditions d'apprentissage ont été différentes et que leur input peut avoir été plus ou moins restreint.

En fait, seules les structures les plus fréquentes auront été fréquemment entendues par les locuteurs de tous les groupes. Les structures moyennement fréquentes n'auront été entendues fréquemment que par le groupe majoritaire. Et les structures plus rares auront été entendues par le groupe majoritaire de façon à être globalement acquises bien qu'il y ait quelques lacunes, alors que ces mêmes structures auront été trop peu entendues par les autres groupes pour être acquises.

C'est la raison pour laquelle il faut considérer non seulement la fréquence d'une structure, mais aussi sa fréquence par rapport à une concurrente, et surtout sa présence dans l'input d'un locuteur.

## **2. Origines, catégories et explications des liaisons**

Nous venons de voir quelques théories générales concernant le MBU, le rôle de la fréquence et de la productivité des structures dans l'acquisition d'une L1 et d'une L2. Voyons maintenant quelques études importantes sur les liaisons.

Les études portant sur les liaisons dans le cadre du MBU sont rares et celles qui abordent les liaisons dans cette perspective se limitent principalement, à celles de Bybee (2001, 2002, 2005), Boula de Mareüil et Adda-Decker (2002) et Boula de Mareüil, Adda-Decker et Gendner (2003). Par contre, d'autres chercheurs ont étudié les liaisons, en

mentionnant brièvement le rôle de la fréquence dans la production de liaisons, sans pour autant s'y attarder.

Pour mieux comprendre le MBU et pourquoi nous avons choisi ce cadre de recherche, nous établirons un lien entre les résultats des études précédentes et le MBU.

La prononciation des liaisons, tout comme l'évolution de la langue en général, est déterminée par plusieurs facteurs dont l'évolution « naturelle » de la langue qui a tendance à la simplifier et à la régulariser. Comme nous venons de le voir, la fréquence peut être une force motrice dans la réduction et la grammaticalisation, mais aussi une force conservatrice (Diessel, 2007). Les formes irrégulières moins fréquentes sont régularisées par un effet d'analogie tandis que les formes irrégulières ou moins productives plus fréquentes sont préservées de l'effet d'analogie.

Selon Morin et Kaye (1982 : 291), c'est surtout le facteur historique qui influence la conservation et donc la production de liaisons. Selon Mallet (2008 : 24), les règles du bon usage peuvent aller à l'encontre de la langue qu'utilise le peuple, parfois par souci de pureté ou de beauté de la langue, parfois pour s'en démarquer, pour que seule l'élite puisse la connaître et la maîtriser : « La réintroduction des consonnes finales orthographiques (XVII<sup>e</sup> siècle), réalisée de manière plus ou moins anarchique, ne fera qu'accentuer le clivage entre les plus érudits (qui pratiquent le bon usage) et les 'masses populaires'. Comme aiment le rappeler les spécialistes de la pédagogie contemporaine, la graphie ne rend plus compte des usages et l'orthographe n'est plus la langue, elle n'en est plus que 'l'habit' ».

Les changements ont parfois pour but d'éviter des confusions, comme par exemple l'ajout du *g* à *vint* (qui devient *vingt*) pour éviter la confusion avec le passé simple de venir (il *vint*) (Malécot, 1954 : 290). Les règles du bon usage consistent alors soit à conserver des mots ou structures disparus ou sur le point de disparaître, soit à créer de nouvelles règles qui vont à l'encontre de l'évolution naturelle de la langue.

Un autre phénomène qui a joué un rôle important dans l'évolution de la langue orale est paradoxalement, le développement de l'écrit. Alors qu'au départ la langue écrite reproduit l'oral, à partir de l'imprimerie et à partir du moment où les écrits sont plus accessibles, les règles de l'écrit ont tendance au cours des siècles à s'imposer sur l'oral. Les règles du bon usage sont utilisées à l'écrit et l'écrit sert de base à l'oral.

Pour savoir lire correctement, il faut d'une part connaître les règles de phonétique française et leurs exceptions (le plus souvent intuitives pour les francophones), d'autre part avoir une connaissance et une mémoire auditive des mots et séquences de mots pour savoir comment les prononcer correctement ; enfin, il faut connaître la correspondance entre graphie et phonie. Si une personne ne connaît pas très bien les règles phonétiques mais sait prononcer des mots et des phrases, la mémoire phonétique peut compenser les lacunes en connaissances phonétiques et orthographiques.

Toutefois, si une personne ne connaît pas la prononciation d'un mot parce qu'elle ne l'a jamais entendu (ou pas assez souvent) par exemple, le mot *gageüre*, les connaissances phonétiques intuitives et les connaissances des règles orthographiques risquent d'être insuffisantes pour lire le mot correctement.

De même, si un mot ou une séquence de mots a souvent été entendu, mais la graphie est assez éloignée de la phonie du mot, ou bien qu'un mot ou une séquence de mots puisse avoir un homonyme, un lecteur peut ne pas faire la correspondance entre graphie et phonie, ce qui peut entraîner des erreurs de lecture. Par exemple, un sujet peut faire la liaison en utilisant en parole spontanée l'expression : *il a trouvé le pot aux roses*, mais ne pas faire cette même liaison en lisant la phrase écrite, ne reconnaissant pas qu'il s'agit de la même phrase et pensant peut-être que l'expression s'écrivait *le poteau rose* !

Ainsi, pour être capable de lire correctement, rapidement, sans erreurs et sans hésitations, il faut connaître les règles orthographiques (reconnaître les mots à l'écrit), les

règles phonétiques (savoir comment prononcer des mots et des phrases) et être capable de faire le lien entre graphie et phonie, et ce, en temps très limité.

Or, la fréquence joue un rôle primordial tant dans la reconnaissance phonétique que dans la reconnaissance graphique des mots. Plus les mots ont été rencontrés à l'écrit, plus il est facile et rapide de les lire. Plus les mots ont été rencontrés à l'oral, plus il est facile de les prononcer correctement. Précisons que pour les apprenants d'une langue seconde, ce n'est pas parce que des mots sont connus, entendus et répétés souvent qu'ils sont systématiquement faciles à prononcer.

Si les mots ont été souvent rencontrés à l'écrit et à l'oral, ces mots ne devraient créer aucune difficulté pour les locuteurs natifs, et moins de difficultés aux apprenants que les mots moins connus. Inversement, si des mots sont rares à l'écrit et à l'oral, alors ils devraient créer quelques difficultés, demander un temps de réflexion plus long et susciter quelques hésitations, reprises ou erreurs, dépendant de l'expérience personnelle du locuteur.

Enfin, en nous basant sur les études du MBU, ainsi que celles de Bybee (2001, 2002, 2005), nous pensons que, si les mots sont rares à l'écrit et à l'oral et sont des exceptions aux règles phonétiques (h aspiré, liaisons interdites etc.), ces mots seront les plus susceptibles de causer des erreurs de lecture en L1 comme en L2. De même, un faible taux de cooccurrence entre deux mots suscitera un temps de réflexion plus long et des possibilités d'erreurs plus grandes. Quant aux liaisons, elles auront moins de chances d'être produites si l'un ou plusieurs des paramètres suivants sont présents : faible taux d'occurrence de la structure, faible taux d'occurrence du mot1, faible taux d'occurrence du mot2, faible taux de cooccurrence des deux mots, structure complexe, structure peu productive regroupant un nombre restreint d'items. Plusieurs facteurs rentrent en jeu en même temps.

Résumons quelques-uns des principes du MBU dont les éléments ci-dessous sont expliqués en plus de détail dans Tomasello (2000a et 2000b) et Ellis (2002) :

1. Pour qu'une structure, une séquence ou un item soit acquis, ils doivent avoir été entendus et utilisés un grand nombre de fois. La fréquence est l'un des facteurs les plus importants dans l'acquisition et la maîtrise d'une structure et d'un item.
2. Les apprenants passent par des étapes :
  - a. D'abord ils apprennent des items isolés, au coup par coup.
  - b. Puis ils « créent des classes et catégories », des règles, et représentations abstraites ;
  - c. Une fois qu'ils ont entendu et utilisé des items et des structures un grand nombre de fois, ils maîtrisent ces structures et font peu d'erreurs (ou pas du tout). Les erreurs portent sur des items isolés et non sur les principes de la structure.
  - d. S'ils maîtrisent les structures, ils sont capables de généraliser des règles et de produire des énoncés corrects qu'ils n'ont jamais entendus à partir d'éléments connus.
3. Si les locuteurs ont peu entendu des séquences ou des structures, ils vont utiliser les structures les plus fréquentes comme référence et faire des erreurs de régularisation, surgénéralisation ou analogie ou au moins hésiter et se reprendre.
4. Les structures ou les items plus rares demandent un temps de réflexion et de traitement plus long et modifient le traitement habituel d'une séquence. Lors de la lecture, un sujet peut hésiter, s'arrêter, se reprendre, ne pas faire une liaison, faire une liaison sans enchaînement (voir l'étude du projet PFC), ou faire une erreur notamment dans le traitement de la consonne de liaison.
5. Les erreurs sont représentatives du processus d'acquisition, de la fréquence des items ou des structures entendues dans l'environnement langagier des sujets et des systèmes maîtrisés par les locuteurs.
6. On retrouve le même système d'acquisition et les mêmes phénomènes dans toutes les catégories linguistiques (sémantique, morphologie, syntaxe, phonétique).



7. La fréquence d'un type de structure détermine les possibilités créatives ou la productivité d'une construction (Bybee : 2001, 2002, 2005). Plus une structure regroupe d'items, plus le locuteur a appris des exemples de cette structure et mieux il la maîtrise. Des structures contenant plus d'items sont plus « productives » que celles qui en contiennent moins.
8. Un sujet peut maîtriser une structure sans en maîtriser tous les items (exceptions), et réciproquement : un sujet peut maîtriser quelques items sans pour autant maîtriser toute la structure correspondante (c'est le cas notamment pour les h aspirés).
9. Une fois les structures apprises, la fréquence ne joue plus un rôle aussi important. C'est surtout en début d'acquisition que la fréquence joue un rôle majeur, lors de la création de représentations abstraites.
10. Les compétences langagières (compréhension et production) sont basées sur une analyse statistique qui détermine les probabilités de production de structures et séquences les plus fréquentes (voir par exemple Saffran et al, 1996 ; Juliano et Tanenhaus, 1993).
11. Les éléments sont codifiés dans un système hiérarchique. Les éléments les plus entendus se trouvent en haut de la hiérarchie et sont le plus facilement accessibles d'après le « Modèle de la Concurrence » (*the Competition Model* de Mac Whinney, 1978, 1982, 1987).
12. La fréquence conditionne l'évolution de tous les aspects de la langue. C'est une force motrice dans la réduction, la régularisation et la grammaticalisation de structures et séquences mais peut aussi être une force conservatrice, comme avec les verbes irréguliers qui ont résisté aux changements et aux régularisations. (Diessel, 2007)
13. La fréquence favorise le développement de tronçons préfabriqués. Les enfants comme les adultes considèrent des séquences comme faisant partie d'un tout. (Dugua, 2005 et 2006). Une grande partie de la langue est constituée de séquences préfabriquées (Erman et Warren, 2000).

14. Plus une structure est complexe, plus la fréquence joue un rôle important : la fréquence compense la complexité et les difficultés des apprenants à acquérir cette structure. Plus un item est fréquent, plus l'irrégularité qui lui est reliée est susceptible d'être acquise.
15. Si une structure ou une séquence est complexe, regroupe peu d'items, et que ces items sont rares, elle va causer plus de difficultés aux locuteurs que celles qui contiennent des items plus fréquents ou celles qui regroupent plus d'items.

Ainsi, en gardant le MBU à l'esprit, nous analyserons brièvement quelques faits concernant l'évolution de la prononciation du français en rapport avec les liaisons, pour mieux en comprendre les mécanismes et les difficultés qui en découlent. Puis, nous verrons les classifications des liaisons, les problèmes liés à ces classifications, les différentes approches tentant d'expliquer les liaisons et les productions de liaisons d'après les études les plus connues sur le rôle de la fréquence et enfin le lien entre les liaisons et le MBU. Nous verrons pourquoi le MBU permet d'expliquer certains phénomènes que la phonologie et la syntaxe ont des difficultés à expliquer isolément ou même considérés conjointement.

## **2.1. Les liaisons et l'évolution de la langue**

Les liaisons présentent des caractéristiques très particulières qui sont dues en partie à l'histoire de la langue. Les consonnes finales étaient toutes prononcées jusqu'au XI<sup>e</sup> siècle. À partir des XI<sup>e</sup> et XII<sup>e</sup> siècles elles se sont progressivement amuïes et n'étaient d'abord plus prononcées que devant une voyelle (enchaînement) et ensuite avec des liaisons. Certaines liaisons ont survécu à l'évolution de la langue et ont gardé leur ancienne prononciation. Comme l'explique Delattre (1955 : 42) : « Du point de vue historique, la liaison résulte généralement de la survivance d'enchaînements de consonnes finales de l'ancien français. Ces consonnes se sont conservées dans la liaison entre deux mots étroitement unis, à cause du haut

degré de syllabation ouverte qui caractérisait la prononciation de la fin du Moyen Français. »

Au cours des siècles, seules les liaisons les plus fréquentes ont survécu, mais elles ont parfois gardé certains traits typiques d'une époque, les rendant aujourd'hui particulièrement irrégulières. Parmi les vestiges de l'ancien français, on trouve trois prononciations de consonnes finales pour les numéros 6, 9 et 10 (Malécot, 1954 : 287), selon qu'ils se trouvent devant une voyelle ou une consonne. Huchon (2002) discute de ce phénomène :

La chute des consonnes finales intervient dans la langue populaire, alors que la langue savante tend à les conserver à la pause et en liaisons (témoigne encore de ce phénomène la prononciation moderne des nombres six, huit : huit enfants, ils sont huit s'opposent ainsi à huit livres).  
Les consonnes implosives (c'est-à-dire en fin de syllabe et devant une autre consonne) disparaissent de la prononciation au XI<sup>e</sup> -XII<sup>e</sup> siècle, entraînant un système de syllabes ouvertes (consonne + voyelle) [...] (Huchon, 2002 : 72)

Huchon (2002) explique la complexité des traités de prononciation, gardiens du bon usage :

[...] La prononciation au XVI<sup>e</sup> siècle fait l'objet de traités de prononciation, rédigés en latin pour les étrangers, de considérations dans un certain nombre de grammaires et de systèmes d'écritures phonétiques, témoignages toutefois souvent contradictoires en raison de la multiplicité des usages, des particularités régionales fortement marquées, des différences entre langue populaire et langue savante, entre langue orale et langue écrite [...] (Huchon, 2002 : 152)

Puisque la prononciation des consonnes finales (avec ou sans liaison) était aléatoire et dépendait des grammairiens de l'époque et du bon usage dicté par les personnes influentes à la cour ou d'autres personnes de pouvoir, la langue évolue selon des règles parfois arbitraires et indépendantes de l'évolution des langues comme le rapporte Huchon (2002) :

[...] La prononciation des consonnes à la pause dans la déclamation soutenue explique les règles de la versification classique qui interdisent de faire rimer des mots avec des consonnes finales différentes. Toutes les consonnes finales sont sourdes ; ainsi *David* rime avec *vit*. Mais il y a, dans la graphie, par souci étymologique, établissement de sonores au lieu de sourdes qui sont, toutefois, conservées dans la prononciation en liaison (*grant* de *grandis* devenant *grand*, mais prononcé avec un [t] dans *grand enfant* ; *sanc* de *sanguinem* devenant *sang*, mais prononcé avec un [k], prononciation usitée en français moderne dans le *sang impur* de *La Marseillaise*). L'influence de l'orthographe sur la prononciation est importante ; [...] (Huchon, 2002 : 154-155)

Autrement dit, la règle phonétique qui consiste à prononcer le d de *grand enfant* comme un [t], (et plus rarement le g comme un [k]) est pour le moins complexe, et ce, d'autant plus pour des apprenants du français L2. Si de telles règles n'ont pas été apprises formellement, la seule façon de savoir prononcer les liaisons dans ces contextes est de les

avoir entendues de nombreuses fois pour en avoir une mémoire auditive.

En fait, les adjectifs sont rarement placés avant le nom, et la liaison entre un adjectif antéposé se terminant par un d + un nom se réduit à quelques items. De fait, ni la logique, ni l'analogie ne permettent de savoir qu'il faut prononcer le d comme un [t], puisque dans d'autres cas (sauf avec *second*), le d ne se prononce pas systématiquement [t], et parfois ne se prononce pas du tout en liaison comme par exemple dans *blond enfant* (on dirait plutôt *un enfant blond*).

Vu le rapport de plus en plus opaque entre la graphie et la phonie, il y a au XVI<sup>e</sup> siècle un débat entre ceux qui veulent que l'écriture reflète l'oral et ceux qui veulent garder la tradition écrite. Huchon (2002) remarque que ce débat est toujours d'actualité, ce qui a provoqué des réformes de l'orthographe notamment en 1740, 1835, 1878, 1935 et 1990.<sup>13</sup>

[...] Dès le milieu du (XVI<sup>e</sup>) siècle s'opposent les partisans d'une orthographe « miroir de la parole », dont on éliminerait toutes les lettres superflues ou polyvalentes (par exemple *c* et *g* prononcés différemment dans *car* et *cent*, dans *gare* et *gent*) ou équivalentes (par exemple, *x*, *s* ou *z* en finale et *en* et *an*), et les tenants de l'usage qui mettent en avant la nécessité de marquer l'origine, les dérivations et de distinguer des homonymes. [...] De fait, les réformateurs comme Meigret ou Peletier n'obtiennent pas gain de cause, l'usage étant celui des ateliers d'imprimerie où l'on s'est refusé à toute transformation radicale ou systématique. [...] (Huchon, 2002 : 157)

La langue évolue au cours des siècles, surtout avec l'usage du peuple, mais les tenants de l'écriture (imprimeurs, grammairiens, écrivains...) en empêchent une progression trop rapide. Grammairiens, phonéticiens et autres pédagogues écrivent des ouvrages contenant les règles normatives de l'élite, tels que Vaugelas (1647), Passy (1892), Martinon (1913), Grammont (1914), Fouché (1959) ou Grevisse (1969) pour ne citer que les plus célèbres ainsi que l'Académie Française. Ceux-ci déterminent le bon usage et permettent aux lecteurs d'éviter des erreurs. À cause de ces règles normatives et à cause d'autres facteurs (historiques), la langue orale comprend de nombreuses exceptions. La langue écrite est de moins en moins le reflet de la langue orale, notamment en ce qui concerne les liaisons.

De plus, des prononciations différentes se manifestent selon les classes

---

<sup>13</sup> Pour plus de détails, voir [http://www.academie-francaise.fr/langue/rectifications\\_1990.pdf](http://www.academie-francaise.fr/langue/rectifications_1990.pdf) p.4

sociales, comme le suggère Huchon (2002) :

L'usage normé distingue les liaisons obligatoires des liaisons facultatives, liaisons dont l'usage s'était particulièrement développé au cours du XIX<sup>e</sup> siècle dans les classes moyennes. Les liaisons facultatives sont de moins en moins observées. L'absence de liaison peut devenir un marqueur sociolinguistique (prononciation populaire de petit à petit sans le [t] de liaison). L'instabilité des liaisons entraîne un certain nombre d'hypercorrections qui amènent des liaisons fautives : cuirs (fausses liaisons en [t]) ; velours (fausses liaisons en [z]) ; pataquès (substitution d'une consonne à une autre). (Huchon, 2002 : 253)

La langue orale est liée à l'écrit et l'écrit se base sur des règles de grammaire plus ou moins arbitraires. En témoigne l'évolution particulière du h aspiré vs le h muet dont discute Huchon (2002 : 152-153) : « Le h aspiré, articulé dans les mots d'origine francique, tend à s'amuïr, ce qui est une "faute énorme" pour Dolet, commune chez les Auvergnats, les Provençaux, les Gascons et les provinces de langue d'Oc ».

Certaines règles ou structures sont très difficiles à apprendre pour tous les francophones (sans parler de ceux qui ne sont pas francophones) parce que les items qui en constituent la classe sont rares et peu entendus par les locuteurs. De plus, si la structure va à l'encontre de concurrentes plus fréquentes, la tendance des locuteurs sera de traiter, par effet d'analogie, cette structure plus rare et plus complexe comme une structure plus fréquente, et de produire des séquences qui ne correspondent pas aux normes prescrites (et donc de faire des erreurs).

C'est le cas par exemple avec les h aspirés. Selon les explications traditionnelles, on utilise un h muet pour les mots d'origine grecque ou latine et un h aspiré pour les mots d'origine germanique ou autre origine étrangère. Il faut soit que les locuteurs connaissent ou devinent l'origine étymologique et connaissent la règle pour savoir que ces mots ont un h aspiré et qu'on ne doit donc faire ni la liaison, ni l'élision, soit qu'ils apprennent le statut phonologique de la syllabe initiale du mot en même temps qu'ils apprennent l'élément lexical.

Toutefois, comme le précise Southworth (1970 : 67), on trouve bien souvent des mots d'origine grecque ou latine avec des h aspirés, si bien que même des connaissances étymologiques ne peuvent pas aider les locuteurs. Les raisons pour lesquelles certains mots

ont un h aspiré en dépit de leur origine étymologique, seraient pour produire un effet « expressif, descriptif, onomatopéique, euphonique ».

D'ailleurs, toujours selon Southworth (1970), certains mots n'avaient pas de h pour commencer, tels que *huile* venant de *oleum* ou *huître*, venant de *ostrea* ; au contraire, certains mots ont tout simplement perdu leur h (*assassin* venant de l'arabe *hacchâchi* passé ensuite dans l'italien *assassino*, ou bien le mot *auberge* qui vient de l'allemand *haribergon*).

Le h aspiré est parfois utilisé pour distinguer certains mots d'autres mots (*haut/eau*). Toutefois certaines règles proviennent d'erreurs d'hypercorrection ou de lectures ou de conseils de personnes influentes de l'époque où les changements phonétiques ont eu lieu (*huile*). Enfin, pour certains h aspirés, il n'y a pas de raison du tout, comme dans les mots « hurler, hennir, hérisson, haleter... » puisque ces mots ne sont pas d'origine germanique, et ils n'ont pas d'homonymes créant des confusions potentielles (Southworth, 1970 : 69).

Par ailleurs, certains mots n'ont pas de liaisons ou d'élisions dans certains cas, mais en ont dans d'autres : les huit (pas de liaison) mais dix-huit, vingt-huit (liaison). On retrouve ce même phénomène avec les autres numéraux qui n'ont pas de liaison non plus, bien qu'ils n'aient pas de h (*un, onze*). Il existe également un « flottement » dans l'usage de certaines semi-consonnes telles que *la ouate* mais un tampon *d'ouate* ou avec certains mots ayant un h aspiré. Dans le mot *hiatus*, la lettre h peut être un h aspiré ou un h muet, c'est-à-dire que pour ce mot les deux h sont acceptables.

Enfin, on trouve des cas pour lesquels certains francophones produisent une liaison alors que les règles l'interdisent, par exemple dans la séquence *les haricots*. Comme le h présente des difficultés dans le choix de réalisation de la liaison et de l'élision, une tendance naturelle des locuteurs francophones est de simplifier la langue et ses règles. Puisque la majorité des h sont muets, la tendance logique serait de progressivement prononcer les h aspirés comme des h muets. C'est effectivement le cas notamment avec *les haricots*. Il existe

alors un décalage entre l'usage du peuple et la norme.

D'un côté, l'évolution de la langue (la langue du peuple) tend à en simplifier les règles en éliminant les exceptions, donc à faire une liaison, d'autre part la norme stipule le contraire, comme on le voit avec les explications de l'Académie Française :

Le haricot ou l'haricot ?

Le h de haricot est « aspiré », c'est-à-dire qu'il interdit la liaison, impose que ce mot soit prononcé disjoint de celui qui le précède, au singulier comme au pluriel. On écrit et dit : le haricot, non l'haricot ; un beau haricot, non un bel haricot. Tous les dictionnaires indiquent par un signe conventionnel quels h (généralement d'origine germanique) sont aspirés et quels h (généralement d'origine gréco-latine) ne le sont pas. Pour certains mots, l'usage est indécis. Ce n'est pas le cas de haricot : la liaison est incontestablement une faute. La rumeur selon laquelle il serait aujourd'hui d'usage et admis que l'on fasse cette liaison a été colportée par un journal largement diffusé dans les établissements scolaires, L'Actu (n°8 du jeudi 3 septembre 1998, p.7), qui n'a pas jugé bon de publier de rectificatif<sup>14</sup>.

Les définitions de Malécot (1975 : 173) rendent bien compte de l'évolution de la langue et de ce phénomène : « Forbidden liaisons involve word-final consonants that have now completely disappeared before vowels, that optional liaisons involve those that are now in the process of disappearing, and that the required ones involve those that have not yet begun to erode, but where the dividing lines are, no one can say. »

La langue en perpétuel changement est tiraillée entre l'évolution naturelle et la norme qui résiste souvent à cette évolution et qui est très présente et puissante en France. À l'école, les enfants apprennent le français de la norme (écrite), la grammaire, l'orthographe et la littérature. La norme de l'oral, c'est la norme de l'écrit, quand c'est possible, et sinon du français recherché (pour les aspects phonétiques), comme l'explique Mallet (2008 : 20) : « [...] la scolarité obligatoire couvre au minimum dix à douze des premières années de la vie et [...] l'apprentissage explicite de l'orthographe y est central. Quotidiennement exercé et renforcé, l'enseignement de la grammaire scolaire, avec ses manipulations de listes de mots, de règles et d'exceptions et ses raisonnements paradigmatiques exemplaires, est au cœur du dispositif pédagogique. »

Selon Pagliano et Laks (2008 : 6), la langue écrite peut même aller à l'encontre de la

---

<sup>14</sup> <http://www.academie-francaise.fr/langue/questions.html#haricot>

langue orale : « [...] l'évolution phonétique régulière peut être troublée par l'influence de l'orthographe. La langue ne s'apprend pas seulement par l'oreille, mais aussi, et surtout de nos jours, par les yeux, ce qui a pour résultat qu'en voulant parler comme on écrit ou épelle, on finit par prononcer les lettres purement étymologiques. »

Mais si la langue enseignée à l'école est surtout la langue écrite, comment un enfant apprend-il les spécificités de la langue orale ? La langue orale qu'apprend un locuteur natif est celle de son entourage. C'est par mimétisme (ou comme le dit Thomas, 1998, « par osmose »), par imitation qu'un locuteur a une intuition de la manière de prononcer certains phonèmes, mots ou tronçons.

Plus on entend de séquences et de structures, plus on est capable d'adapter sa production à des situations d'énonciation particulière et d'utiliser correctement la langue orale recherchée, la langue du pouvoir, la langue d'une population instruite. Gadet (1989 : 52) précise que plus les gens sont instruits, plus ils peuvent adapter leur discours et plus leur discours présente de la variation : « On peut dire, en suivant Encrevé, que la liaison est « un phénomène sociolinguistique inversé », car ce sont les locuteurs les plus scolarisés qui présentent sur elle le plus large système de variation. »

Cette langue comporte en plus des règles de l'écrit (en syntaxe et morphologie) certaines règles (telles que les règles des liaisons) que seules les personnes les plus éduquées connaissent et maîtrisent. Les politiciens et les personnes de pouvoir utilisent, pour la plupart, la prononciation du français recherché et les spécialistes de la parole utilisent même des structures uniques à leur groupe linguistique telles que les liaisons non-enchaînées (Encrevé, 1983, 1988). Par exemple, dans l'étude de Malécot (1975), ce sont les administrateurs qui font le plus de liaisons.

Enfin, les nombreuses parodies sur la prononciation d'hommes politiques et de célébrités, se moquant des personnes qui parlent « mal » incitent les locuteurs à se forcer à



utiliser un « bon » français, c'est-à-dire le français dicté par la norme, même si celle-ci va parfois à l'encontre de ce qui leur paraît naturel.

Les locuteurs n'utilisent plus leur intuition pour des structures plus rares, car ils ne les ont pas assez entendues pour les maîtriser et les connaître avec certitude, mais ils utilisent leur intuition de ce qui leur paraît être plus approprié d'après ce qu'ils ont entendu plus rarement de la part de personnes éduquées, par exemple à la télévision. C'est ainsi qu'ils peuvent faire des cuirs, velours et pataqués, dans des situations d'hypercorrection.

Il résulte de tout ceci que l'évolution linguistique associée aux facteurs étymologiques, historiques et sociopolitiques a participé à créer la langue telle que nous la connaissons aujourd'hui, dont les règles phonétiques sont pour le moins difficiles à décrire. Ce qui constitue les principaux motifs de sa prononciation, c'est l'usage du peuple, qui est lui-même influencé par la fréquence des mots et des séquences. Or, l'usage est parfois en désaccord avec les règles normatives et ces règles ne correspondent pas toujours à la réalité linguistique.

Afin de mieux comprendre toute la complexité du système de liaisons, nous allons maintenant évaluer les descriptions, catégorisations et explications des liaisons, que nous comparerons ensuite aux productions de francophones.

## **2.2. Les catégories de liaisons**

Les règles sur les liaisons, qu'elles aient été établies pour les francophones (l'Académie Française) ou pour les apprenants du français L2 ou FLE posent certaines questions auxquelles nous allons répondre :

1. Quelles sont les définitions, catégories de liaisons ?
2. Sont-elles en accord les unes avec les autres ? Sinon, pourquoi ?
3. Pourquoi les catégorisations sont-elles problématiques ?
4. Pourquoi les explications « classiques » ne permettent-elles pas vraiment de rendre compte des phénomènes apparaissant avec les liaisons ?

5. Pourquoi les règles, explications, descriptions et théories « classiques » ne permettent-elles pas de prédire les productions de liaisons ?

Si des règles tentent d'expliquer le fonctionnement de l'oral, elles vont forcément se heurter à plusieurs problèmes en se limitant à une description des catégories syntaxiques. Si elles généralisent des règles syntaxiques, cela implique que de nombreux cas n'y appartiendront pas, ou le contraire, que des structures devraient répondre à certaines conditions alors que ce n'est pas le cas. C'est ce que remarque Mastromonaco (2000) :

In order for liaison to occur, both the phonological and syntactic conditions on liaison must be satisfied. If only one of these conditions is satisfied, liaison will either not occur where it should, or will occur where it should not. For example, if only the syntactic conditions for liaison are respected, the floating consonant will be realized between the determiner and noun as in *les garçons* [lezgaRsɔ̃], even though the phonological context does not allow it. Another example of this is a floating consonant that is realized before h-aspiré words, as in *les haricots* [lezaRiko]. If only the phonological conditions are respected, liaison will occur in *ils veulent aller à la plage*, or *l'enfant adorable*, despite the fact that the syntactic context does not allow it (with an infinitive or singular noun). (Mastromonaco, 2000 : 281)

Précisons que la liaison avec un infinitif est possible mais rare. Nous sommes d'accord avec l'auteure que, dans le contexte donné, la liaison aurait moins de chances d'être réalisée, vu qu'il y a une liaison facultative juste avant, dans *veulent aller* et que le groupe de mots est un peu trop long pour permettre une liaison (ce dont nous reparlerons plus tard).

D'autre part, dire que le contexte [adj. + nom] est une liaison obligatoire, comme le précisent l'Académie Française, Delattre (1951) et Encrevé (1988), est problématique, parce que cela ne prend en compte que le côté syntaxique. Si l'on fait effectivement la liaison avec certains adjectifs qualificatifs antéposés tels que *petit* ou *gros*, cette liaison est rarement réalisée à 100% avec tous les adjectifs. On ne peut pas dire que si la liaison n'est pas produite, c'est une erreur.

La liaison est encore plus facultative si les mots sont peu fréquents (*dernier étage*, *léger ennui* dans notre étude) et elle est pratiquement interdite si le taux de cooccurrence des

mots est très faible (*blond enfant* dans Sampson, 2001). Autrement dit, on ne peut pas dire que la liaison avec les adjectifs antéposés est obligatoire, puisque la liaison dans ce contexte est par exemple réalisée à 72.5% dans l'étude de Boula de Mareuil, Adda-Decker et Gendner (2003 : 1331).

D'un autre côté, il serait « faux » ou tout au moins « étrange » de ne pas faire une liaison avec certains adjectifs fréquents tels que *gros*, *grand* ou *petit*. En fait, plusieurs facteurs semblent déterminer la production de liaison, dont la fréquence de cooccurrence des mots. Morin (2005 : 11) considère qu'on fait la liaison dans : « D'innombrables constructions, parfois stéréotypées, telles que : Un ancien n-administrateur, le prochain n-album, d'un commun n-accord, le premier r-étage, dans son dernier r-article, [...] un parfait t-inconnu [...], un dangereux z-individu, [...] un heureux z-événement [...]. »

Toujours dans l'étude de Boula de Mareuil, Adda-Decker et Gendner (2003 : 1331), les liaisons sont réalisées à plus de 90% pour certains items particuliers de certaines classes, sans que cela ne soit le cas pour la catégorie entière. C'est le cas notamment pour certaines prépositions (*en*, *dans*) certaines conjonctions (*quand*) et certains adverbes (*très*). Il est donc problématique d'expliquer cette situation et d'en tirer des règles : pour certains items d'une même catégorie syntaxique les liaisons sont obligatoires, alors qu'elles sont facultatives pour les autres.

En outre, si les règles tentent d'établir des listes plus précises, elles courent le risque d'avoir de très longues listes, avec de nombreuses exceptions, comme c'est le cas par exemple avec les innombrables règles et exceptions de Fouché (1959). Ceci risque en fin de compte d'être contreproductif, car peu de personnes voudraient les lire et encore moins seraient capables de les retenir.

Pour qu'une règle soit représentative de la réalité linguistique, elle doit prendre en compte au minimum les facteurs syntaxique et phonétique, mais considérer la fréquence et la

longueur des mots, ce dont nous allons discuter plus loin. Ceux qui tentent de donner des règles uniquement syntaxiques se heurtent à des problèmes inévitables. Nous prendrons comme exemple les catégorisations de Delattre (1951, 1966) qui nous semblent les plus représentatives et les plus utilisées, notamment par Encrevé (1988) et que l'on retrouve également dans les explications de l'Académie Française.

De nos jours, le Conseil supérieur de la langue française, ainsi que l'Académie Française en France et le Conseil de la langue française du Québec, dictent les règles du bon usage du français. D'après l'Académie Française, les règles sur les liaisons<sup>15</sup> sont les suivantes:

En français, la liaison peut apparaître entre un mot qui se termine par une consonne et un mot qui commence par une voyelle ou un h non aspiré (voir aussi l'article Le haricot), si ces deux mots ne sont séparés par aucune ponctuation ni par aucune pause orale. Selon les cas, elle est obligatoire, facultative ou interdite. Les noms propres sont également soumis à la liaison.

La liaison est obligatoire :

- entre le déterminant et le nom : des(z) amis, tout(t) homme ;
- entre l'adjectif antéposé et le nom : un(n) ancien(n) usage ; ainsi on dira un savant(t)aveugle si aveugle est un nom, mais un savant aveugle si savant est le nom ;
- entre le pronom (sujet ou objet) et le verbe : ils(z)aiment, on(n)aime, ils vous(z)aiment, ils(z)y vont, (...)
- entre est et le mot qui suit, dans des formes impersonnelles ou dans la forme présentative : il est(t)évident qu'il viendra ; c'est(t) à voir ;
- entre l'adverbe et le mot unis étroitement : trop(p) étroit ; bien(n)aise ;
- entre la plupart des prépositions monosyllabiques et le mot qui suit : dans(z) une heure ;
- dans la plupart des mots composés et locutions : un pot(t)-au-feu, mot(t) à mot, de temps(z) en temps.

Elle ne se pratique pas :

- après la conjonction et : un fils et une fille ;
- après la consonne finale d'un nom au singulier : un temps idéal, un nez épaté ;
- après le s intérieur dans les locutions nominales au pluriel : des moulins à vent ;
- après la finale -es de la 2e personne du singulier de l'indicatif présent et du subjonctif présent : tu portes un habit vert ; Il faut que tu lui écrives un poème. On fera en revanche la liaison lors de la lecture de vers ;
- après les mots terminés en -rt en -rs, sauf s'ils sont suivis de il, elle, on ou s'il s'agit du t de l'adverbe fort ou du s de toujours : de part en part, tu pars à huit heures (mais : quand dort-(t)on ? quand sort-(t)elle ?) ;
- devant un, oui, onze et les mots étrangers commençant par y : des oui ;
- devant les noms de lettres de l'alphabet : des i, des a.

Dans le reste des cas, on peut choisir de faire ou non la liaison mais celle-ci est plutôt la marque d'un langage soutenu. On distingue par ailleurs deux types de fautes de liaison :

- le cuir qui consiste à faire une liaison en t à la place d'une liaison en z, et plus généralement à effectuer à mauvais escient une liaison en t : Il s'est mis(t) au travail ; J'ai cru(t) apercevoir un écureuil ;
- le velours qui consiste à faire une liaison en z à la place d'une liaison en t, et plus généralement à effectuer à mauvais escient une liaison en z : vingt(z) euros ; les dix-huit(z) ouvrages ; Il est venu aujourd'hui(z) encore ;

Ces deux types de liaisons fautives sont aussi appelés des pataquès.

Par extension, on désigne par pataquès, cuir ou velours toute liaison fautive, quelle qu'elle soit.

---

<sup>15</sup> <http://www.academie-francaise.fr/langue/questions.html#liaisons> p.18-19

Ces catégorisations sont globalement similaires à celles de Delattre (1951) :

#### **Liaison obligatoire**

article + les\_enfants, un\_enfant, aux\_amis  
adjectif + ces\_amis, deux\_amis, pls\_amis  
pro personnel + vous\_êtes, vous\_y\_êtes, nous\_en  
vbe + pro pers allez\_y, vient\_elle  
est (impers) + c'est\_évident, il est\_interdit  
préposition + dans\_un an, en\_hiver  
adv monosyll + très\_utile, bien\_étrange

#### **Liaison facultative**

nom pluriel + plats\_exquis, mots\_à te  
pro non pers + pls\_écoutent, dont\_elle  
pro pers post + amusons-nous\_un peu  
verbe + suis imberbe, fait\_appel  
préposition + depuis\_un an, avant\_une  
adv polysyll + jamais\_à l'heure, tjs\_ici  
conj monosyll + mais\_alors, puis\_on ira  
X + invariable maisons\_et, bons\_ou

Delattre (1955 : 45) précise que les liaisons dépendent grandement de leur « degré d'étroitesse ». Il précise aussi que les liaisons sont obligatoires entre une préposition et un adverbe « monosyllabique » et prend déjà en compte plusieurs facteurs. Toutefois, de telles catégorisations sont encore incomplètes car les liaisons ne sont pas réalisées de la même façon selon les prépositions et les conjonctions, même si on les sépare en « monosyllabiques » et « polysyllabiques ».

Prenons quelques cas particuliers pour analyser la variabilité des résultats selon certaines catégorisations. D'après les règles de Delattre (1951) et de l'Académie Française, la liaison est obligatoire avec *est* alors que d'après Ågren (1973) cette liaison est une liaison facultative fréquente puisque la liaison de [est + part. passé] se fait dans 98.6% des cas et dans 97% des cas pour *est* en général. Durand et Lyche (2008 : 46) d'après les données du corpus PFC concluent que la liaison avec *est* n'est réalisée que dans 30% des cas.

Un autre exemple est celui de l'adverbe *trop*. Pour Valdman (1993 : 168): « [...] la liaison est obligatoire dans *trop*, mais, précisément parce que ce mot contient /p/ la liaison a tendance à devenir facultative. » D'après les règles de l'Académie Française, elle est obligatoire. Pour Léon (1966 : 129) cette liaison est facultative peu fréquente (ou en voie de disparition) ». Pour Malécot (1975 : 175), elle est « rare en conversation et optionnelle en élocution ». Ainsi, une même liaison est interprétée de façon très différente selon les chercheurs, les époques et les corpus.

Les manuels de phonétique ne s'accordent ni sur les règles ou catégories de liaisons obligatoires, ni sur celles des liaisons facultatives. Comme le constatent Peyrollaz et de Tovar (1954 : 177) : « Dans cette question des liaisons, on ne peut formuler de règle *absolue*, puisque, comme on l'a dit plus haut, l'usage n'est pas fixé. »

Alors que les règles concernant des structures « simples » telles que [adj. + nom], ou [c'est +...] sont déjà sujettes à controverse, les règles et structures plus compliquées telles que les liaisons interdites avec notamment les h aspirés, représentent un enjeu de taille tant dans la description des règles que dans leur acquisition.

La question des h aspirés/vs h muets - et des liaisons interdites en général - est extrêmement complexe. D'abord, la liste des mots commençant par un h aspiré est longue. Grevisse (1980 : 59) en dénombre 190, et Fouché (1959 : 252-258) 677. D'autre part, les mots de ces listes exhaustives sont des mots rares, voire très rares (*hampe, hauban, huard...*). En revanche, si seuls les mots rares avaient un h aspiré, cela simplifierait la tâche, mais ce n'est pas le cas. De nombreux mots rares ont également un h muet (*hanséatique, héraldique, héroïde...*). Si les mots commençant par un h (h aspiré ou h muet) sont des mots rares, donc peu entendus, les locuteurs ont peu de chance de savoir s'il faut faire l'élision et la liaison.

À la liste « classique » de mots commençant par un h aspiré, s'ajoutent d'autres listes, les autres mots avec un h aspiré avec lesquels on fait une liaison dans les noms communs et les noms propres. Fouché (1959 : 258-265) en fournit huit pages d'explications et de descriptions. Les règles diffèrent, selon que les noms propres sont français ou francisés, des noms composés, des appellations nobiliaires, des anciens noms propres du Moyen-âge, des mots de provinces ou nations étrangères. Par contre, de nombreux noms propres pourtant d'origine étrangère ont un h muet. Fouché (1959 : 263) ne cite que les 72 plus « fréquents ». Pour certains noms de famille étrangers portés par des Français, l'h est aspiré lorsque les mots sont monosyllabiques, mais l'h est muet quand ils sont polysyllabiques (Fouché, 1959 : 264).

En dehors des listes de mots rares, certains mots plus fréquents ont un h aspiré avec un morphème et un h muet avec un autre morphème, bien que les deux mots aient le même sens et la même étymologie, comme c'est le cas avec *le héros* vs *l'héroïne*. Le mot *héros* a un h aspiré pour le distinguer de son possible homonyme en cas de liaison (*les héros* vs *les zéros*). Il en est de même avec d'autres mots tels que *les hauts* vs *les eaux*.

De plus, il existe un flottement pour de nombreux mots, notamment pour les mots *hiatus* et *h*. Par exemple, selon le Petit Robert (2002), il est également acceptable de les prononcer avec un h aspiré ou un h muet : ['aʃ ; aʃ] (p.1234) et [ʒatys ; 'ʒatys] (p. 1266).

Enfin, même si les locuteurs connaissent les mots les plus fréquents, s'ils les ont entendus prononcés avec une liaison, comme dans *les\_haricots*, alors qu'il n'en faut pas, cela fausse leurs représentations abstraites, car dans leur lexique, ces mots se prononcent avec une liaison. Il apparaît donc que la liaison avec les h aspirés est extrêmement problématique. C'est l'un des éléments que nous testerons dans notre étude.

Boula de Mareüil, Adda-Decker et Gendner (2003 :1330), grâce à l'étude du corpus BREF contenant 91 126 occurrences de contextes de liaisons parmi lesquels 40 940 liaisons ont été observées (soit 45%), font une comparaison des règles classiques des liaisons et de la réalité linguistique observée dans leur corpus. Les résultats obtenus ne correspondent pas entièrement aux définitions. D'après ces auteurs, pour qu'une liaison puisse être considérée comme « obligatoire », elle devrait être réalisée au moins à 95%. Par contre, leur corpus ne sépare pas les cas de déterminants suivis de contextes empêchant les liaisons tels que h aspirés, semi-consonnes, etc. De plus, il y a toujours en parole spontanée des cas d'hésitations, reprises, des erreurs de performance, etc., si bien que leurs résultats devraient être considérés avec précaution, surtout si on les compare aux résultats d'autres études.

Comme on peut le voir dans le tableau 1.1., la liaison dite « obligatoire » n'est vraiment obligatoire dans ce corpus qu'après les déterminants, entre un verbe et un pronom et

après *quand*. Elle est très fréquente après certaines prépositions monosyllabiques. Par contre, elle ne semble pas être obligatoire dans tous les contextes mentionnés par l'Académie Française, notamment entre un adjectif et un nom (72.5%).

structure syntaxique	Nb occ.	% réalisé pour cette structure	% de tous contextes réalisés
déterminant +	15272	95.3	35.6
auxiliaire ou verbe +	6997	81.0	13.6
pronom clitique +	4534	83.3	9.2
prép. monosyllabique +	4269	90.5	9.4
adverbe monosyllabique pas +	2570	70.7	4.4
adjectif + nom	1733	72.5	3.1
verbe + pronom	1081	99.2	2.6
quand +	168	94.6	0.4

**Tableau 1.1. % L et structures syntaxiques. Boula de Mareüil, Adda-Decker et Gendner (2003 : 1331).**

Après un déterminant, la liaison est obligatoire car elle a été réalisée dans plus de 95% des cas. C'est la structure la plus fréquente représentant plus d'un tiers de toutes les liaisons observées. Si l'on considère ce contexte plus en détail (tableau 1.2.), on voit qu'à l'intérieur d'une même catégorie, on trouve des sous-catégories qui semblent dépendre, elles aussi, de la fréquence de ces structures. Par exemple, le contexte [dét. + nom] semble être plus fréquent que le contexte [dét + forme verbale] (de type adjectif participial), ce qui correspond aux pourcentages de liaisons réalisées.

structure syntaxique	Nb occ.	% réalisé pour cette structure	exemples
tous les dét. +	15272	95.3	
déterminant + pronom	376	98.7	les uns
déterminant + nom	14077	96.6	son image
déterminant + adjectif	713	95.5	un obscur
déterminant + forme verbale (nom basé sur un participe passé)	33	60.6	des élus
déterminant + numéral	57	4.2	les un
déterminant + autre	16	-	

**Tableau 1.2. % de liaisons réalisées des déterminants. Boula de Mareüil et Adda-Decker (2003 : 1332).**

Les liaisons qui n'ont pas été réalisées après un déterminant ne l'étaient pas parce que le mot2 avait un h aspiré ou appartenait à la catégorie des exceptions. En d'autres termes la liaison est généralement réalisée après un déterminant, tant que le contexte phonétique le



permet (sauf si le mot2 contient un h aspiré, une semi-consonne, une autre exception, ou tout autre phénomène empêchant la liaison).

Dans la même étude de Boula de Mareüil, Adda-Decker et Gendner (2003), la plupart des liaisons facultatives sont réalisées à moins de 50%, mais certaines (considérées comme fréquentes ou très fréquentes) sont réalisées en grandes proportions. Certaines liaisons dites interdites ont été réalisées entre 1.1% dans le contexte [et + ...], et 10.4% dans le contexte [nom commun sing. + nom]. Si une liaison est censée être « interdite » d'après la norme, mais que dans la réalité elle est quand même réalisée à 10%, on peut considérer à juste titre que la liaison n'est pas interdite, mais plutôt facultative rare.

C'est la raison pour laquelle il n'y a pas l'unanimité entre la norme des manuels de phonétique, de l'Académie Française, des définitions et catégorisations des grammairiens et des linguistes et des dictionnaires. Par exemple, dire que la liaison est obligatoire après un adverbe monosyllabique est partiellement juste puisqu'elle est réalisée à 94% avec *très* dans le corpus BREF de Boula de Mareüil, Adda-Decker et Gendner (2003 :1330) et à 96.55% dans le corpus PFC (Mallet : 281).

Mais dire que la liaison est obligatoire après un adverbe monosyllabique est problématique puisqu'avec l'adverbe *pas* elle n'est réalisée qu'à 40% dans le corpus BREF et à 1.36% dans le corpus PFC. Certains adverbes monosyllabiques ont une liaison obligatoire, d'autres une liaison facultative fréquente et d'autres enfin ont une liaison facultative rare. Plusieurs facteurs rentrent en jeu, si bien qu'il est difficile de formuler des règles phonétiques et de déterminer si les liaisons sont obligatoires, facultatives, interdites ou même flottantes. En fait, on peut tout au mieux parler de « tendances ».

De telles définitions bien qu'apparemment normatives ont tout de même leur place, car elles proposent des directives précieuses aux locuteurs francophones qui veulent apprendre les règles du « bon parler » et éviter « les fautes ». Ces règles sont d'autant plus

importantes pour des apprenants du français L2 ou FLE. Le problème n'est donc pas d'avoir des règles, mais plutôt d'avoir des règles qui sont un peu plus proches de la réalité linguistique francophone.

De nombreuses règles « classiques » ne représentent pas la réalité linguistique pour les raisons suivantes :

1. Elles sont principalement syntaxiques et ne prennent pas en compte les autres dimensions linguistiques (phonématiques, lexicales, etc.).
2. Elles sont normatives et non descriptives.
3. Elles se basent sur les normes d'un français qui ne représente pas le français parlé actuel, tel qu'il est utilisé par la majorité francophone.
4. Elles ne prennent pas en compte l'effet de la fréquence.
5. Elles ne considèrent pas les effets cumulés de plusieurs facteurs.

Une solution consisterait donc à prendre en considération plusieurs dimensions linguistiques, tout en gardant des règles simples et relativement courtes. Par exemple, une explication ou catégorisation un peu plus juste en ce qui concerne la liaison des adjectifs antéposés serait :

- la liaison est très fréquente avec les adjectifs les plus fréquents tels que *petit, gros* ; etc. et lorsque le taux de cooccurrence du mot1 et du mot2 est élevé : *petit enfant...*
- La liaison est de moins en moins fréquente lorsque les mots et le taux de cooccurrence des mots sont plus rares. Par exemple, dans la séquence : *un hautain enfant*, la liaison est rare et semble même impossible.
- Plus le mot1 est court, plus il est susceptible de lier.
- La liaison est fréquente avec les phonèmes /z/, /n/ et /t/ (*petits enfants, bon été, petit ours*).
- La liaison est rare avec le phonème /R/ (*premier étage, dernier effort*).

- La liaison est très rare avec le phonème /p/.

Il faut toutefois admettre que de telles descriptions risquent d'être un peu plus longues et aussi un peu plus compliquées que celles qui existent aujourd'hui mais au moins auront-elles le mérite de se rapprocher davantage de la réalité linguistique.

### **2.3. Les explications « classiques »**

Comme nous venons de le voir, les descriptions et catégorisations classiques ne correspondent pas complètement à la réalité linguistique et permettent difficilement de prédire les réalisations de liaisons. Nous mentionnerons ici brièvement les diverses théories sur les liaisons, surtout pour en déterminer les limites et pour expliquer le cadre théorique de notre thèse : une approche lexicale multidimensionnelle basée sur le MBU.

#### **Une explication historique**

Nous avons abordé brièvement certains éléments historiques et avons vu que des événements marquants ont contribué à l'évolution de la langue, en préservant certaines formes, en éliminant d'autres et en créant ou changeant des règles telles que celles des h muets et h aspirés. Certaines liaisons se faisaient il y a quelques décennies et ne se font plus aujourd'hui.

Il semble que les francophones fassent de moins en moins de liaisons (facultatives) comme en témoignent les études de Thomas (1998 : 546) et Malécot (1977) par rapport à Delattre (1951), d'Ashby (2003) par rapport à l'étude d'Ashby (1981) au même endroit et dans les mêmes conditions, ainsi que l'étude de Smith (1995-6) par rapport à celle d'Ågren (1973). Les comparaisons entre les productions de liaisons d'il y a plusieurs décennies comparées aux productions plus récentes soulignent un déclin dans plusieurs catégories de liaisons, surtout pour les liaisons facultatives. Les liaisons obligatoires semblent plus stables.

Comme le précise Delattre (1955 : 48) : « Le facteur historique répond de variations qui ont pu se maintenir au cours de plusieurs siècles, indépendamment du style, de l'union ou des conditions phonétiques subséquentes. » Toutefois, des études assez anciennes (Delattre, 1947) font état de pataquès qui ne sont toujours pas acceptés aujourd'hui, bien qu'ils continuent d'être produits dans des contextes similaires. Le facteur historique comprend aussi la norme qu'imposent certains groupes (politiciens, spécialistes du langage, éditeurs, etc.). Toutefois, ce facteur à lui seul ne permet pas d'expliquer certains phénomènes tels que l'usage du peuple, les erreurs que font les enfants et les adultes et la variabilité des liaisons.

### **Une explication syntaxique**

Comme nous venons de le voir, les explications les plus fréquentes sur les liaisons sont syntaxiques. En effet, le facteur syntaxique est primordial : si l'on compare les études de corpus (Malécot, 1975 ; Ågren, 1973 ; Ahmad, 1993, par exemple), on observe qu'il existe une hiérarchie des liaisons d'un point de vue syntaxique. On peut établir un continuum des liaisons les plus souvent produites, au moins souvent produites. Les classifications de Boula de Mareüil, Adda-Decker et Gendner (2003 : 1331) sont elles aussi principalement syntaxiques (sauf pour les classifications en termes de prépositions ou adverbes monosyllabiques).

Par contre, comme l'indique De Jong (1988), le lien syntaxique n'explique pas à lui seul la variabilité et ne peut en aucun cas prédire quelle liaison sera effectuée dans quelle situation. En outre, tous les adjectifs, tous les auxiliaires et toutes les prépositions ne sont pas traités de la même façon.

Il est difficile de trouver une description syntaxique unifiée, dans la mesure où les exigences syntaxiques seules ne peuvent expliquer pourquoi une liaison se fait plus avec une personne, qu'avec une autre, plus avec un auxiliaire qu'avec un autre, ou plus avec une

préposition plutôt qu'une autre. Par exemple dans l'étude d'Ågren (1973), la liaison avec *est* est produite à 98.7% tandis qu'avec *suis* elle n'est produite qu'à 47%. Plus précisément, la liaison avec [est + part. passé] se fait à 98.6% alors qu'avec [suis + part. passé], elle ne se fait qu'à 57%. Selon Bybee (2001), cela peut s'expliquer par le fait que la 3<sup>e</sup> personne du singulier est plus fréquente que les autres personnes.

Morin et Kaye (1982 : 312-313) expliquent que les généralisations de Selkirk (1972) concernant l'effet du trait morphologique d'inflexion ne sont pas satisfaisantes pour les raisons suivantes : quelquefois l'hypothèse de Selkirk est trop forte et prédit qu'il n'y aura pas de liaisons alors que celles-ci sont possibles. En outre, Mastromonaco (281) précise que la théorie de Selkirk est trop faible et prédit parfois une liaison qui est en fait impossible, si l'on ne prend en compte que le facteur syntaxique. Fougeron et al (2002) en viennent à la même conclusion, selon laquelle la théorie X' ne permet ni d'expliquer ni de prédire la liaison :

La c-commande, essentiellement utilisée pour les relations anaphoriques et les règles d'accord, est définie chez Reinhart ('81), comme : A c-commande B, si A ne domine pas B, B ne domine pas A, et le premier nœud branchant dominant A domine également B. [...] Dans le cas de « *enfants ont* » [...], la condition de c-commande n'est pas réalisée : le nœud NP, qui est le premier nœud branchant dominant le nœud N « *enfants* », ne domine pas l'auxiliaire « *ont* » en position T', et la liaison n'a pas lieu. Le contexte « *conduits en* » montre que la c-commande est une condition nécessaire mais pas suffisante pour la liaison : le lieu c-commande le lié, mais la liaison est intuitivement peu probable, voir interdite. (Fougeron et al., 2002 : 3)

Ågren (1973) et De Jong (1994) ont eux aussi remarqué que, dans un même contexte syntaxique, les items ne recevaient pas le même traitement. Par exemple, De Jong souligne que la liaison est réalisée dans 20% des cas avec *était/étaient* mais seulement dans 5.3% des cas avec *étais*.

On ne peut donc pas seulement parler de façon générale d'adjectifs, pronoms, prépositions, conjonctions ou auxiliaires, mais plutôt de préposition monosyllabique fréquente ayant tel ou tel phonème, d'auxiliaire être à la 3<sup>e</sup> personne du singulier, etc. pour rendre compte des liaisons les plus fréquemment réalisées (Bybee, 2001, 2005). Or, la simple

précision *monosyllabique* ou *polysyllabique* signale qu'une explication uniquement syntaxique est insuffisante et propose un accord avec une approche partiellement lexicale.

### **Une explication morphologique**

Pour Laks (2005), la liaison est surtout déterminée par les frontières de mot et les consonnes finales graphiques. Les notions de fréquence d'usage reflètent les catégories principales toujours réalisées qui sont morphologiques : le marquage du nombre et de la personne; c'est-à-dire que plus les mots sont proches syntaxiquement et sémantiquement, plus on fait la liaison.

En effet, le /z/ de liaison est plus fréquemment réalisé entre des déterminants et des noms, des adjectifs et des noms et avec des pronoms et des verbes qu'avec des prépositions ou des formes verbales fléchies. De même, le facteur morphologique peut expliquer pourquoi on peut faire la liaison dans *des nez aquilins* ou *des bois immenses*, alors qu'on ne la fait pas au singulier avec *un nez aquilin* ou *un bois immense*.

Le développement langagier des enfants suggère que les consonnes de liaison sont traitées comme des « préfixes du nom ». Un nom comme *enfant* aurait des formes fléchies telles que « *enfant, zenfants, tenfant, nenfant* » (Morin, 2003), ce qui correspond à des formes partiellement morphologiques (z pour le pluriel, n pour le singulier, etc.). En cours d'acquisition, les enfants agglutinent une consonne avec le mot suivant « le n-âne, le petit z-enfant » car ils n'ont pas encore acquis les contraintes, les restrictions « qui régissent ces combinaisons ».

On trouve d'ailleurs le même type d'erreurs chez les adultes avec des groupes nominaux pluriels, que ce soit avec les adjectifs numéraux (*quatre-z-yeux*) ou dans les erreurs à distance (*des chefs d'état-z-africains*) (Tranel, 1981 : 216), parce que les locuteurs adjoignent le morphème du pluriel à des items qui ne le permettent pas.

Toutefois, une explication uniquement morphologique ne peut pas expliquer de nombreux phénomènes, tels que le fait que les locuteurs du projet PFC ont fait la liaison dans 100% des cas avec *sommes* mais seulement dans 20.83% des cas avec *avons*, bien qu'il s'agisse du même marqueur de la 1<sup>e</sup> personne du pluriel. De même comment expliquer que la liaison a été produite dans 100% des cas après *sous*, contre 94.97% avec *dans*, 83.33% avec *chez*, et 0% avec *vers*, alors qu'il s'agit de la même catégorie de prépositions monosyllabiques? (Mallet, 277).

Enfin, l'un des marqueurs morphologiques les plus représentatifs est le s du pluriel. Or, le corpus PFC fait état de plusieurs cas où le s du pluriel n'est pas prononcé, comme on s'y attendrait. Par exemple, Durand et Lyche (2008 : 45) remarquent que le s de l'adjectif *premiers* n'est pas toujours prononcé, et le s de *petits* a été prononcé plusieurs fois [t] au lieu de [z] comme par exemple dans *petits hôtels*. De même, la liaison avec le s de *ils* n'est pas toujours réalisée en français canadien, comme l'observent Durand et Lyche (2008 : 57) : « Canadian varieties also demand further attention: e.g. *ils* before a vowel-initial verb can be realized without a plural [z] (e.g. [ij\_õ] for *ils ont*); »

Une explication uniquement morphologique ne permet donc pas de rendre compte de tous ces phénomènes.

### **Une explication phonologique et prosodique**

Une explication phonologique permet d'expliquer plusieurs phénomènes. De nombreuses règles concernant les liaisons sont phonétiques et prosodiques. Par exemple, la liaison se fait principalement avec les phonèmes /z/, /t/ et /n/, plus rarement avec /R/ et de façon exceptionnelle avec /p/.

Une langue à syllabation ouverte telle que le français a tendance à éviter l'hiatus. En effet, on trouve en français de nombreux phénomènes qui ont pour but de l'éviter : les liaisons

et enchaînements, l'élision, l'ajout de consonnes épenthétiques et euphoniques et autres stratégies d'évitement du hiatus tels que *cet arbre, mon amie, arrive-t-il, etc.*

Delattre (1955) aborde le facteur prosodique avec la longueur des éléments à lier, l'intonation de la phrase et l'accent d'insistance. En effet, l'accent d'insistance et le coup de glotte empêchent la liaison ou tout au moins en réduisent les réalisations. Delattre (1955 : 46-47) aborde également le facteur phonétique :

A. La liaison se fait plus fréquemment après voyelle qu'après consonne. Des faits historiques [fe zistorik] lie plus que: Des heures historiques [œr zistorik]. Deux raisons se présentent: d'abord il est plus aisé d'articuler une consonne que deux; ensuite le besoin de liaison se fait moins sentir lorsqu'il y a déjà enchaînement: [œristorik], puisqu'il n'y a pas de hiatus à éviter. Dans les paires d'exemples ci-dessous, la phrase de gauche lie plus fréquemment que celle de droite.

Des noms amusants-Des contes amusants  
Des volets ouverts-Des violettes ouvertes [...]

B. La liaison se fait plus facilement après une consonne qu'après deux.

Des contes amusants-Des nombres amusants  
Des heures historiques-Des actes historiques [...]

C. Peut-être aussi la liaison est-elle moins fréquente quand le hiatus à éviter est entre deux voyelles de timbre différent: Nous avons été, Tu as été, que lorsqu'il est entre deux voyelles de même timbre: Vous avez été.

D. Enfin si le participe présent résiste moins à la liaison que l'infinitif, c'est sans doute en partie qu'il se présente plus de liaisons en t (ou s) que de liaisons en r. Le facteur est ici psychologique autant que phonétique.

(Delattre, 1955 : 46-47)

Les liaisons dépendent aussi d'autres phénomènes phonétiques tels que la production de schwa ou d'assimilations. Par exemple si *je suis* est prononcé /ʃyi/, il y a peu de chances pour que la liaison soit ensuite produite.

Durand et Lyche (2008 : 52) concluent toutefois que la liaison ne peut pas être réduite au simple facteur phonologique :

A final complication is that different consonants do not appear to have the same linking capabilities. It seems to us that among the three main liaison contenders (t, z, n), the plosive [t] allows non linking and glottal stop insertion more easily than the fricative [z] and, in turn, that the latter is a better pseudo-coda than [n]. We have neither heard nor seen reported pronunciations such as *certain[n]* [ \_]atout, *bien[n]* [ \_]utile.

In the same way, we conjecture that in corpora of elevated speech one can contrast [r] (a strong forward linker if variable liaison is realized) with [p] (a frequent non linker in attested variable liaisons). [...] the whole gamut of hesitations combined with evidence from experimental phonetics and a careful consideration of morphophonological facts demonstrate that liaison consonants cannot be treated as pure and simple onsets of W2, which would be the naïve interpretation of the PFC codings providing 'liaison enchaînée' in approximately 99.5% of the cases and apportioning the rest to speech errors.

(Durand et Lyche, 2008 : 52)



D'un point de vue uniquement phonologique, ni l'approche générativiste linéaire, avec la règle de troncation de Schane (1965), ni les règles de liaison de Dell (1970, 1973), ni les théories sur l'épenthèse (Klausenburger, 1974, 1976 ; Tranel, 1981 ; Morin et Kaye, 1982), ni l'approche multilinéaire de Goldsmith (1976) ou Clements et Keyser (1981), ou les théories mentionnées par Encrevé (1988) sur les consonnes flottantes (entre autres) ne permettent de répondre de façon définitive à certaines questions. (Pour plus d'explications sur les théories phonétiques, consulter par exemple Mallet, 2008 : 107-130, ou Mastro Monaco, 2000 : 1-77).

Les théories phonétiques ne permettent pas de répondre aux questions suivantes :

1. Pourquoi les liaisons peuvent-elles être obligatoires, facultatives ou interdites dans un même contexte phonétique (dont discute Sampson, 2001) ?  
*ancien élève* liaison réalisée la plupart du temps (considérée comme obligatoire)  
*vain espoir* la liaison n'est pas réalisée
2. Comment expliquer que des prépositions, adjectifs, ou verbes ayant la même consonne finale aient un taux de réalisation différent ? Par exemple, la liaison est réalisée à 43.9% avec *est*, à 30.9% avec *peut*, à 19.2% avec *sont*, etc. (Mallet, 229)
3. Comment expliquer les erreurs de liaisons de type *quatre-z-yeux*, *chemins de fer-z-anglais*, etc. puisqu'il n'y a pas de consonne flottante à la fin de ces mots ?
4. Comment expliquer que les erreurs enfantines du type *un lours* (au lieu de *un-n-ours*), sont similaires à celles des adultes du type *il va-t-en cours* (Dugua 2005, 2006) ?
5. La contrainte anti-hiatus explique difficilement pourquoi certaines liaisons ne se font pas et entraînent un hiatus qui aurait pu être évité (*il est / arrivé*, *je suis / arrivé*).
6. Pourquoi prononcer une consonne de liaison suivie d'une consonne (*vingt et quand*) si la contrainte est que le mot2 commence par une voyelle ?
7. Pourquoi faire une liaison sans enchaînement et pourquoi faire une liaison avec un s, si l'on peut faire un enchaînement, comme dans *toujours arrivé* (Morin, 2005) ?

Nous voyons donc qu'une explication uniquement phonologique ne permet pas d'expliquer tous ces phénomènes ni de prédire les réalisations de liaisons.

### **Une explication orthographique**

La liaison consiste à prononcer, dans certaines conditions, une consonne latente écrite. C'est la raison pour laquelle les liaisons sont réalisées dans une plus grande proportion en lecture, et encore plus en lecture de vers. C'est notamment parce que les consonnes écrites sous les yeux du lecteur incitent à les prononcer davantage qu'en parole spontanée où on ne voit pas ces consonnes.

En parole spontanée, la mémoire auditive est suffisante pour la prononcer correctement les liaisons les plus fréquentes, la plupart des liaisons obligatoires. en effet, un jeune enfant de 5 ans qui ne sait pas encore lire et des personnes analphabètes produisent des liaisons qui n'ont aucun rapport avec l'écriture.

Toutefois, il est indispensable de connaître l'orthographe des mots, pour savoir quand et comment prononcer une liaison qui n'est pas fréquente dans la conversation de tous les jours (liaisons facultatives plus rares). Il faut connaître et reconnaître les formes écrites et être capable d'en faire le lien avec les formes orales. Plus une personne a l'habitude de lire, plus elle rencontre et reconnaît des structures et des items, plus elle peut en produire en parole spontanée.

On observe ce phénomène en particulier dans les discours de personnes médiatiques qui produisent des consonnes de liaison non enchaînées, dans des discours écrits oralisés ou la prononciation consiste en grande partie à prononcer les mots tels qu'ils sont écrits. Une telle prononciation donne l'impression que le locuteur est érudit, car il parle « comme un livre ».

Précisons que la fréquence des formes écrites est différente de celle des formes orales. Une séquence peut être fréquente à l'orale, mais peu fréquente à l'écrit, et *vice versa*. C'est ce

qu'on peut constater en analysant des séquences d'après une banque de données de l'écrit telle que *Frantext* (élaborée à partir de livres), et en les comparant avec une banque de données de l'oral telle que *Lexique*<sup>16</sup> (élaborée à partir de sous-titres de films). Ainsi, l'écrit est important pour reconnaître des mots et séquences qui sont fréquentes à l'écrit.

Par contre, une règle seulement orthographique ne permet pas d'expliquer les erreurs de liaison consistant à insérer une consonne épenthétique qui ne pourrait jamais se trouver dans ce contexte, par exemple : *j'ai failli-t-être, ou quatre-z-yeux*. Tout francophone qui maîtrise l'orthographe du français sait que *quatre* ne prend jamais de *s*. Pourquoi fait-on la liaison dans *des bois immenses* mais pas dans *un bois immense*, alors qu'il s'agit pourtant de la même consonne écrite. Pourquoi fait-on la liaison dans 90% des cas avec *petites*, mais seulement à 50% avec *grandes*? Une explication orthographique ne permet donc pas de rendre compte de certaines erreurs et de la variation des liaisons.

### **Une explication stylistique**

Le registre de langue est un facteur déterminant dans la production de liaisons. Plus le registre est élevé, plus on fait de liaisons. C'est pourquoi les discours présentés dans les médias par les politiciens et les journalistes sont ceux qui contiennent le plus de liaisons facultatives (Encrevé, 1988).

En plus du registre, le ton est aussi un facteur important. Selon que le ton est ironique, solennel ou humoristique, il n'y aura pas le même nombre de liaisons. D'ailleurs, une même personne, dans un même discours, peut produire plus ou moins de liaisons facultatives, selon les variations de ton dans un même discours (Ashby, 1981).

Selon les termes de Malécot (1975) certaines liaisons sont plus « prestigieuses » que d'autres. Delattre (1947 : 149) remarque qu'on fait de plus en plus de liaisons selon les

---

<sup>16</sup> <http://www.lexique.org/moteur/RechCoOc.php>

situations d'énonciations suivantes : la conversation familière, la conversation soignée, la conférence et enfin, la récitation des vers. Parmi les facteurs qui favorisent la production de liaisons se trouvent les conditions d'énonciation qui vont déterminer le registre de langue, le sujet de la conversation, le ton, etc. mais aussi les caractéristiques personnelles du locuteur : l'âge, le niveau social, l'éducation, le nombre d'années d'étude, l'origine géographique, etc.

Comme le précise Gadet (1989 : 52) : « La liaison constitue un indicateur sociolinguistique explicitement très fort, qui à lui seul permet de classer socialement un locuteur. [...] »

Malécot (1975 : 168) considère plusieurs facteurs sociolinguistiques tels que la situation d'énonciation et certaines caractéristiques du locuteur : « [...] some corresponding to content (banalities, narration, explanation, reflection) and some pertaining to the attitude or social posture of the speaker (degree of formality, attitude toward interlocutor, etc.) ».

Enfin, si le facteur stylistique permet d'expliquer pourquoi les locuteurs font plus de liaisons dans certaines conditions d'énonciation, il ne peut pas expliquer pourquoi un même locuteur, dans un même discours et dans une même situation énonciative, fera des liaisons dans certains cas mais pas dans d'autres.

### **Une explication lexicale**

L'influence de la fréquence lexicale sur la production de liaisons a été considérée dans plusieurs études, dans lesquelles les liaisons semblent plus souvent réalisées avec des mots fréquents que rares (Ågren, 1973). Bybee (2001) explique que dans un même contexte phonétique et syntaxique, seul l'effet de fréquence permet d'expliquer les différences de productions de liaisons.

De plus, la longueur des mots à lier joue un rôle important, rôle connu depuis longtemps. En témoigne la classification de Delattre (1951) qui précise que la liaison est

obligatoire avec un adverbe monosyllabique mais facultative avec un adverbe polysyllabique. De même, Encrevé (1988) remarque qu'avec des lieurs monosyllabiques la liaison est réalisée à 77%, alors qu'elle n'est réalisée qu'à 29% pour les lieurs polysyllabiques. Ågren (1973), DeJong (1991) et Morin et Kaye (1982) font le même constat.

Toutefois, là aussi, une explication uniquement lexicale ne peut expliquer certains phénomènes. La longueur des mots et la fréquence des mots ne peuvent expliquer à elles seules pourquoi certains mots courts et fréquents ont un faible taux de liaison alors que certains mots moins fréquents ont un taux plus élevé. Par exemple, l'adverbe *pas* est court et très fréquent, mais sa liaison est réalisée dans une proportion beaucoup plus faible que des adverbes moins fréquents. Dans le corpus PFC, la liaison avec l'adverbe *très* a été réalisée dans 96.55% des cas, contre 64.11% avec l'adverbe *plus* et 1.36% avec l'adverbe *pas*. Une explication uniquement lexicale et fréquentielle ne permet pas d'expliquer les différences de réalisations entre les liaisons.

## 2.4. Le cumul des facteurs

Chaque facteur considéré isolément ne permet pas d'expliquer ou de prédire les liaisons de façon fiable. Au mieux, chaque facteur peut établir des « tendances ». Comme le dit Mallet (2008 : 293) : « Cela implique que soit l'on a affaire à de simples tendances, soit qu'il faut revoir les catégories et les contextes proposés (cf *très* et *bien*). On ne peut en tout cas pas tenir une généralisation sur toutes les prépositions +, toutes les conjonctions +, etc. ».

Considérer seulement deux facteurs demeure insuffisant. En se concentrant sur les facteurs syntaxiques et phonétiques sans prendre en compte le rôle de la fréquence lexicale, on encourt le risque auquel fait face Sampson (2001 : 254-255). Celui-ci essaie d'expliquer pourquoi dans un même contexte phonétique et syntaxique, parfois la voyelle précédant la consonne de liaison est nasale (*mon ami*), parfois orale (*Moyen-âge, plein air*) et parfois la

liaison n'est pas faite du tout comme avec les adjectifs *fin, hautain, lointain, malin, mignon, souverain*. Il en conclut que la prononciation d'items tels que *mon* est différente des autres à cause de son statut de clitique tandis que la prononciation de certaines séquences telles que *plein air* peut être attribuée à une lexicalisation pré-nominale partielle :

Instead, the appropriate interpretation would seem to be a form of partial lexicalization has taken place with these adjectives when they appear in pre-nominal position, resulting in familiar, stereotypical phrases of the type *ancien élève, Moyen-âge, vain espoir, en plein air*. (Sampson, 2001 : 254).

[...] The exceptional cases noted would seem to be due to two factors: for items like *mon*, the maintenance of a distinct liaison pronunciation is attributable to their clitic status; for items like *bon*, their partial lexicalization pre-nominally within stereotypical phrases is decisive. (Sampson, 2001: 255).

Les hypothèses proposées n'expliquent pas vraiment la variation, et ne permettent pas de prévoir les réalisations de liaisons, dans un même contexte syntaxique (avec les adjectifs antéposés). L'auteur ne parvient pas à expliquer les différences de réalisations, et ne parvient pas à proposer des généralisations applicables aux autres adjectifs, et encore moins aux autres catégories syntaxiques (noms + adjectifs, prépositions et adverbes).

Par contre, la fréquence de cooccurrence des mots et la fréquence du mot1 et du mot2 (que l'on peut trouver dans la banque de données de *Lexique*<sup>17</sup>), permettent une explication plus fiable qui peut mieux prédire les réalisations de liaisons. C'est le taux de fréquence des groupes de mots qui en détermine la prononciation. Dans un même contexte syntaxique, plus la fréquence des mots est élevée, plus la liaison va être réalisée et inversement, moins la fréquence est élevée, moins la liaison va être réalisée.

Si l'on classe les séquences étudiées par Sampson (2001) contenant un adjectif selon leur fréquence, en utilisant *Lexique* (d'après un million de mots), (tableau 1.4.) et que l'on analyse la production de liaison de séquences très fréquentes (liaisons obligatoires avec *mon, ton*), (tableau 1.3.), on voit que lorsque le taux de cooccurrence est très élevé, la liaison est toujours produite par tous les francophones majoritaires. Lorsque le taux de cooccurrence est

---

<sup>17</sup> <http://www.lexique.org/moteur/RechCoOc.php>

assez élevé, la liaison est une liaison facultative fréquente et si les mots sont peu fréquents et ont un taux de cooccurrence très faible (voire inexistant) la liaison est rare.

On voit aussi que le taux de fréquence du mot2 permet d'expliquer un taux de réalisation légèrement supérieur (4 liaisons réalisées sur 11 avec *souverain air*).

Séquences avec adjectif avec nasale	taux de cooccurrence <i>Lexique</i>	réalisations et conditions de réalisations
mon ami	+ de 1000	Liaison systématique avec voyelle nasale/orale
bon ami	321	Liaison systématique avec voyelle orale
Moyen-âge	21	Liaison systématique avec voyelle orale
ancien élève	19	Liaison facultative avec voyelle orale

**Tableau 1.3. Liaisons avec adjectifs avec nasales, fréquence et réalisation**

Séquences étudiées Sampson (2001)	taux de cooc. <i>Lexique</i>	Fréquence mot1 Lexique (film)	Fréquence mot2 Lexique (film)	réalisations et conditions de réalisations des liaisons d'après Sampson (2001)
Souverain air	0	1	473	4/11 (dont 2 avec voyelle nasale)
Lointain avenir	0	5	72	3/11 (dont 2 avec voyelle nasale)
Fin observateur	1	2	1	2/11 (dont 1 avec voyelle nasale)
Mignon objet d'art	0	46	26	1/11 avec voyelle nasale
malin espoir	0	33	57	1/11 avec voyelle nasale

**Tableau 1.4. Séquences étudiées par Sampson (2001) et fréquence de cooccurrence d'après *Lexique***

On voit ici que la réalisation des liaisons dépend de la fréquence du mot1 et du mot2 et de la fréquence de cooccurrence des mots. On voit aussi que la fréquence détermine la prononciation des consonnes de liaison. Si les mots sont très fréquents, la consonne de liaison perd sa nasalité (sauf pour *mon*, dont la prononciation peut être nasale ou orale) tandis qu'avec les mots plus rares, les locuteurs de l'étude de Sampson ont prononcé la majorité des consonnes de liaisons avec une voyelle nasale (7 sur 11).

La question de la variabilité et non prévisibilité est au centre des études sur les liaisons. Dans un même contexte syntaxique [adj. + nom], et phonétique (nasale), les résultats sont fortement influencés par d'autres facteurs tels que la fréquence des mots1 et des mots2 et la fréquence de cooccurrence des séquences des mots impliqués. Plusieurs facteurs doivent donc être pris en considération conjointement dont au moins les facteurs syntaxique, phonétique et fréquentiel.

Puisque la mémorisation ne concerne pas seulement des mots isolés mais surtout des séquences, des constructions et des structures, on ne peut pas séparer le facteur lexical du facteur syntaxique. Il y a également une interaction entre la syntaxe et la phonologie (Langacker, 1987 ; Bybee, 1998), mais aussi entre la syntaxe, la phonologie, la sémantique et la fréquence.

Nous allons d'abord évaluer certains facteurs séparément avant de déterminer, dans la discussion à la fin de cette thèse, les facteurs les plus importants dans la production de liaison.

## **2.5. Les liaisons chez les adultes francophones**

Les résultats des études de corpus de français oral varient selon les époques (variation diachronique), les situations d'énonciation, le registre de langue (variation diaphasique), les facteurs relatifs aux sujets tels que leur niveau d'éducation, leur âge, etc. et la variation géographique et régionale (variation diastratique).

On observe de grandes différences (variation diaphasique) entre les personnes utilisant un registre soutenu telles que les politiciens de l'étude d'Encrevé (1988) qui produisent un grand nombre de liaisons et dont un certain nombre de liaisons sont non-enchaînées, et les personnes qui utilisent une langue « ordinaire » (Gadet, 1989 ; Laks, 1983).

La prononciation des sujets est différente (variation diastratique) si les études portent par exemple sur le français parisien (Moisset, 2000), le français tourangeau (Ashby 1981 et 2003), le français méridional (Ranson, 2008) ou le français canadien (Ashby, 1988 ; Ostiguy et Sarrasin, 1985; Walker, 1984 ; Van Ameringen et Cedergren, 1981).

On voit également cette variation diaphasique et stylistique (liaison en lecture ou en conversation) et diastratique (selon les pays et les régions francophones) dans les résultats du projet PFC (Phonologie du Français Contemporain)<sup>18</sup> dont les auteurs Durand, Laks et Lyche

---

<sup>18</sup> [http://www.projet-pfc.net/index.php?option=com\\_wrapperetview=wrapperetItemid=151](http://www.projet-pfc.net/index.php?option=com_wrapperetview=wrapperetItemid=151)



(2002 et 2005) expliquent les principes fondamentaux.

Seules les études récentes de corpus de francophones permettent de rendre compte de la réalité linguistique actuelle. Or, les résultats des études dépendent des méthodes et du contenu de la recherche (syntaxe, phonétique, psycholinguistique, etc.) et des hypothèses que les auteurs cherchent à confirmer ou infirmer. En ce qui concerne les liaisons, des descriptions uniquement syntaxiques (comme celle de l'Académie Française ou celle de Delattre, 1951), phonétiques ou même syntaxiques et phonétiques (Sampson, 2001), permettent difficilement d'expliquer les productions de liaisons et leur variabilité.

Les études de corpus font remarquer que les taux de réalisation des liaisons constituent davantage un continuum dont les limites sont difficiles à préciser que de véritables catégories. C'est la raison pour laquelle les études les plus récentes évitent même d'utiliser des termes tels que *obligatoires (ou catégoriques)*, *facultatives (ou variables)* et *interdites (erratiques, non réalisées)*, et choisissent plutôt de déterminer les pourcentages de liaisons réalisées dans chaque catégorie sans les nommer. Il est en effet plus juste d'utiliser des termes moins catégoriques tels que « le plus souvent réalisée », « réalisées dans un style soutenu » etc. Toutefois, pour faciliter et délimiter les catégories de notre étude, nous conserverons les termes obligatoires, facultatives et interdites, terminologie dont nous reparlerons dans la discussion finale. Nous passerons en revue diverses études et discuterons de quelques exemples pour illustrer la problématique des liaisons.

Tout d'abord, les catégories de liaisons ne sont pas clairement délimitées. Selon Booij et de Jong (1987), les seuls contextes où les liaisons sont réalisées à 100% sont :

1. après un déterminant,
2. après un pronom préverbal,
3. entre un verbe et un pronom (*vas-y, prenez-en*),
4. dans certaines expressions figées (*tout à l'heure*).

Par contre, selon Malécot (1975 : 163-164) aucune liaison n'est réalisée à 100% sauf avec la préposition *en*. La liaison est réalisée

1. avant un nom dans 98% des cas (entre déterminant et nom).
  2. après tous les articles dans 99.5% des cas dont
    - 99.6% avec un nom pluriel
    - 99.1% avec un nom singulier
  2. après un pronom personnel sujet dans 98.6% des cas
  3. après un pronom personnel complément dans 97.1% des cas.
  4. après des pronoms non personnels (autres, plusieurs...) à 50% des cas
  5. après une préposition dans 98.8% des cas, dont
    - 98.8% après une préposition monosyllabique,
    - 100% après *en* et *chez*,
    - 98.8% après *dans* et 85.7% après *sans*.
- Aucune liaison n'a été réalisée après une préposition polysyllabique.
6. après un adjectif dans 93.2% des cas
  7. avec *être* auxiliaire 69.8% et *être* verbe 71.9%

Certains chercheurs utilisent le terme « article » (Howard, 2005) tandis que d'autres utilisent le terme « déterminant » (Malécot, 1975 ou Booij et de Jong, 1987), alors qu'il ne s'agit pas exactement de la même chose. Le terme *article* est, à notre avis, plus précis et surtout plus restreint que le terme *déterminant* (qui comprend certains adjectifs, dont les possessifs, démonstratifs, etc.). Selon que l'on considère tous les déterminants ou seulement les articles, les résultats seront différents.

Malécot (1975 : 167) remarque que sur 7 consonnes de liaison, 99.5% sont produites avec /z/, /t/ et /n/ (consonnes dentales et alvéolaires) et que le taux de réalisation de la liaison n'est pas très élevé dans les groupes figés. Les résultats varient d'une expression à l'autre. Par exemple, la liaison est réalisée à 57% avec *c'est-à-dire* et à 14.8% avec *tout à fait*. Ces résultats sont pour le moins surprenants, car on s'attendrait à un taux de réalisation beaucoup plus élevé pour des liaisons considérées comme obligatoires.

Fougeron et Delais-Roussarie (2004) ont étudié les liaisons et enchaînements de 27 locuteurs en parole lue et en parole spontanée pour un total de 134 841 mots. Le premier corpus a été produit par 17 locuteurs toulousains âgés de 20 à 59 ans. Le second corpus,

SYLSWISS, a été produit par 10 locuteurs suisses âgés de 20 à 30 ans. Les sujets ont lu un texte composé de passages du journal *Le Monde*, sélectionnés pour inclure au moins deux contextes de liaisons possibles par phrase. Les résultats sont les suivants :

L. potentielles	% liaison réal.	Enchaîn. Pot.	% ench. réal.	% l. lecture	% l. conv.
8-10 / 100 mots	39%	7-8 / 100 mots	60%	42%	36%

**Tableau 1.5. Taux de réalisation de la liaison et de l'enchaînement dans deux corpus, d'après Fougeron et Delais-Roussarie (2004 :2)**

Notons que les pourcentages de liaisons réalisées dans cette étude (42% en lecture et 36% en conversation), sont plus bas que ceux d'autres études telles celles de Malécot et d'Ashby. Ces pourcentages plus bas peuvent s'expliquer par le fait que les auteurs de l'étude ont défini comme contextes de liaisons potentielles, les séquences présentant un mot1 se terminant par une consonne et un mot2 commençant par une voyelle.

Cette définition ne considère pas si une liaison est phonétiquement possible ou non (par exemple, s'il y a un h aspiré, une semi-consonne, etc.), si les liaisons sont obligatoires, facultatives ou interdites. C'est aussi le cas dans les études de Boula de Mareüil, Adda-Decker et Gendner (2003). Comme le reconnaissent Fougeron et Delais-Roussarie (2004 : 2), avec cette classification, le nombre de contextes dits « potentiels » est probablement plus bas que le nombre de contextes où la liaison est réellement susceptible de se produire.

Le projet PFC est un projet qui regroupe les productions codées, disponibles en ligne, de 353 locuteurs (au moment de notre étude) venant de plusieurs pays francophones. Il fournit des renseignements précieux sur les réalisations actuelles de liaisons et de schwas dans différents contextes : lecture et conversation. Pour ce projet, le même système de classification a été utilisé. Si un mot1 se termine par une consonne latente et un mot2 commence par une voyelle ou un h, c'est un contexte de liaison potentielle.

Nous avons classé dans le tableau 1.6. les données de l'étude du projet PFC par ordre décroissant de production de liaisons enchaînées dans la lecture. Ces résultats représentent les

productions de 353 locuteurs, dont des locuteurs africains et maghrébins pour lesquels le français n'est pas la langue maternelle. On peut trouver des informations plus détaillées sur les locuteurs sur le site du projet.

L'étude révèle que tous les groupes ont produit plus de liaisons en lecture qu'en parole spontanée, sauf les groupes de Nantes, les Aveyronnais à Paris, et les Africains de Bangui et du Cameroun qui en ont produit légèrement plus en conversation qu'en lecture.

D'autre part, on peut remarquer une grande variabilité entre les groupes, y compris les groupes d'un même pays (régions de France). Pour un même texte lu, les productions de liaison varient de 66.04% pour le groupe de Lacaune à 56.82% pour le groupe de Vendée.

Notons que sur les 35 liaisons « possibles » du texte à lire, se trouvent

- 18 liaisons « obligatoires » *ira-t-il ; grand émoi ; en effet ; jeux olympiques ; son usine ; grand honneur ; très inquiet ; les élections ; les opposants ; D'un autre ; tout est fait ; on en ; en a ; on est jaloux ; Nous avons ; quelques articles ; des activistes ; dans une.*

- 13 liaisons « facultatives » dont

- 3 fréquentes : *est en ; est en revanche ; ont eu tendance ;*

- 10 facultatives rares : *chemises en soie ; pâtes italiennes ; des circuits habituels ; toujours autour ; visites officielles ; provoquer une ; préparent une journée ; plus à ; vraiment une ; trouver au.*

Notons que d'après Gadet (1989 : 54) la liaison entre un nom au pluriel et un adjectif est « encore plus rare » et la liaison après un infinitif est « exceptionnelle ».

- 4 liaisons « interdites » ou extrêmement rares : *Comment, /en plus éviter ; l'opposition / aurait déclaré ; quelques fanatiques / auraient ; il s'est/, en désespoir de cause.*

Selon Gadet (1989 : 53) la liaison entre le sujet au pluriel et le verbe est réservée à la récitation de vers, et est « possible donc, mais très recherchée », donc extrêmement rare.

Cette étude ne sépare pas les résultats selon les liaisons obligatoires, facultatives ou interdites. Toutefois, puisque les pourcentages s'échelonnent entre 66.04% pour le groupe de Lacaune à 56.82%, pour le groupe de Vendée, cela signifie que, globalement, tous les groupes ont produit les liaisons obligatoires, mais que seuls certains groupes ont produit les liaisons facultatives les plus fréquentes tandis que les liaisons interdites n'ont pas été produites. Il est possible de trouver des données plus précises pour chaque locuteur et chaque occurrence, mais ce n'est pas le but de notre étude ici.

Les groupes africains font dans l'ensemble (sauf pour le groupe d'Abidjan) moins de liaisons que leurs pairs européens : entre 57,53% pour le groupe sénégalais et 36.25% pour le groupe camerounais. Alors que le groupe sénégalais réalise les liaisons obligatoires et facultatives les plus fréquentes, le groupe camerounais ne produit pas de liaisons facultatives et approximativement 72% de liaisons obligatoires.

Notons que le seul groupe qui ait un taux de liaisons non-enchaînées élevé (13.75%) qui reflète, à ce stade, autre chose que de simples hésitations, est le groupe camerounais. Le nombre élevé de liaisons non-enchaînées et le nombre réduit de liaisons enchaînées, suggère que ce groupe parle une variété de français dont les normes phonétiques sont différentes du français standard.

Enfin, le groupe de Chlef (Maghreb<sup>19</sup>) a un taux de réalisation nettement inférieur à tous les autres groupes, avec seulement 3.53% de liaisons réalisées, ce qui représente, en moyenne, moins d'une liaison réalisée par personne et par texte, suggérant soit une maîtrise très réduite du système phonétique français, soit la présence d'une variété dialectale dans laquelle les liaisons sont quasi absentes.

Dans le tableau 1.6., nous avons regroupé dans la catégorie « autre », les liaisons incertaines, épenthétiques et autres catégories.

---

<sup>19</sup> Chlef (Maghreb) sont les descriptions et termes utilisés par les auteurs du site

NR= non réalisées      RE : réalisées et enchaînées      NE : réalisées mais non-enchaînées.

	Enquêtes terminées	loc	lecture					convers.			
		nb	% L.	% L.	% L.			% L.	% L.	% L.	
			NR	RE	NE	autre	NR	RE	NE		
17	Lacaune	13	32.5	<b>66.0</b>	0.7	0.7	56.3	<b>41.9</b>	0.4	1.3	
24	Paris Centre ville	12	33.2	<b>65.4</b>	0.7	0.5	60.7	<b>39</b>	0.0	0.2	
19	Lyon	11	33.2	<b>65.1</b>	0.9	0.6	60.4	<b>38.5</b>	0.6	0.3	
23	Ogéville	11	34.0	<b>64.9</b>	0.3	0.7	56.3	<b>43.5</b>	0	0.1	
28	Rodez	9	34.4	<b>64.8</b>	0	0.7	35.6	<b>62.6</b>	0.6	1.0	
21	Nantes	11	33.7	<b>64.7</b>	0.8	0.5	34.0	<b>65.2</b>	0.3	0.3	
20	Marseille	10	35.0	<b>64.3</b>	0.3	0.3	64.8	<b>35</b>	0.0	0.0	
25	Puteaux-Courbevoie	6	28.5	<b>63.4</b>	5.8	2.1	43.6	<b>53.7</b>	1.9	0.6	
13	Douzens	10	36.3	<b>63.3</b>	0	0.3	43.2	<b>56.5</b>	0.2	0	
6	Brécey	11	33.8	<b>63.1</b>	1.6	1.3	58.2	<b>41.2</b>	0	0.5	
30	Toulouse	14	36.1	<b>62.6</b>	0.4	0.6	59.7	<b>39.9</b>	0.1	0.1	
7	Brunoy	10	37.4	<b>61.9</b>	0.6	0.0	52.9	<b>45.7</b>	0.4	0.8	
27	Roanne	9	37.4	<b>61.8</b>	0	0.7	57.0	<b>41.7</b>	0.1	1.0	
1	Abidjan	14	37.9	<b>61.8</b>	0.2	0	55.5	<b>44.4</b>	0	0	
31	Tournai	12	38.0	<b>61.7</b>	0.2	0	55.5	<b>44.2</b>	0.1	0	
5	Biarritz	12	37.2	<b>61.2</b>	1.0	0.5	57.0	<b>41.9</b>	0.5	0.3	
15	Grenoble	12	37.9	<b>61.0</b>	1.0	0	49.1	<b>48.8</b>	0.7	1.2	
2	Aix Marseilles	8	38.5	<b>60.3</b>	0	1.0	69.5	<b>30.3</b>	0	0.0	
12	Domfontais	12	38.7	<b>60.1</b>	1.0	0.0	60.5	<b>38.3</b>	0.1	0.9	
22	Nyon	12	39.0	<b>59.9</b>	0.2	0.7	52.7	<b>46.7</b>	0.1	0.3	
11	Dijon	8	39.3	<b>59.8</b>	0	0.8	61.5	<b>37.2</b>	0.3	0.8	
3	Aveyronnais à Paris	12	37.9	<b>59.6</b>	0.5	1.8	35.3	<b>63.5</b>	0.8	0.2	
18	Liège	12	39.4	<b>59.6</b>	0.5	0.5	48.8	<b>51.0</b>	0	0.1	
16	Ile de la Réunion	9	40.2	<b>59.4</b>	0.3	0	51.2	<b>48.2</b>	0.5	0	
14	Gembloux	12	41.5	<b>57.9</b>	0	0.5	52.3	<b>47.5</b>	0.0	0	
29	Sénégal Dakar	18	42.5	<b>57.4</b>	0	0	45.4	<b>54.5</b>	0	0	
8	Burkina Faso	12	42.9	<b>57.0</b>	0	0	56.3	<b>43.3</b>	0.1	0.1	
32	Vendée	8	41.6	<b>56.8</b>	0.7	0.7	54.6	<b>44.5</b>	0.1	0.7	
4	Bamako	13	48.0	<b>51.5</b>	0.4	0	50.6	<b>49.3</b>	0	0	
26	RCA Bangui	13	51.7	<b>46.4</b>	1.3	0.5	51.2	<b>48.2</b>	0.2	0.2	
9	Cameroun	5	49.3	<b>36.2</b>	13.7	0.6	60.1	<b>37.9</b>	1.1	0.7	
10	Chlef	12	96.4	<b>3.5</b>	0	0	98.7	<b>1.2</b>	0	0	

Tableau 1.6. % de réalisation de la liaison (lecture et conversation) - données du projet PFC en ligne.

Si l'on sépare les résultats par régions (tableau 1.7.), on remarque que mis à part les groupes français de Vendée et de Dijon, qui représentent 8 locuteurs pour chaque groupe, les Français de France métropolitaine produisent dans l'ensemble plus de liaisons que les Suisses et les Belges et les Réunionnais, qui en produisent eux-mêmes plus que les Africains.

<b>France Lorraine</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Ogéville	11	34.0	64.9	0.3	0.7	56.3	43.5	0	0.1
<b>France Bretagne</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Nantes	11	33.7	64.7	0.8	0.5	34.0	65.2	0.3	0.3
<b>France Midi-Pyrénées</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Lacaune	13	32.5	66.0	0.7	0.7	56.3	41.9	0.4	1.3
Rodez	9	34.4	64.8	0	0.7	35.6	62.6	0.6	1.0
Toulouse	14	36.1	62.6	0.4	0.6	59.7	39.9	0.1	0.1
<b>France Languedoc</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Douzens	10	36.3	63.3	0	0.3	43.2	56.5	0.2	0
<b>France Rhône Alpes</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Grenoble	12	37.9	61.0	1.0	0	49.1	48.8	0.7	1.2
Lyon	11	33.2	65.1	0.9	0.6	60.4	38.5	0.6	0.3
Roanne	9	37.4	61.8	0	0.7	57.0	41.7	0.1	1.0
<b>France Ile de France</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Aveyronnais à Paris	12	37.9	59.6	0.5	1.8	35.3	63.5	0.8	0.2
Brunoy	10	37.4	61.9	0.6	0.0	52.9	45.7	0.4	0.8
Paris Centre ville	12	33.2	65.4	0.7	0.5	60.7	39	0.0	0.2
Puteaux-Courbevoie	6	28.5	63.4	5.8	2.1	43.6	53.7	1.9	0.6
<b>France Côte d'Azur</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Aix Marseille	8	38.5	60.3	0	1.0	69.5	30.3	0	0.0
Marseille	10	35.0	64.3	0.3	0.3	64.8	35	0.0	0.0
<b>France Normandie</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Brécey	11	33.8	63.1	1.6	1.3	58.2	41.2	0	0.5
Domfontais	12	38.7	60.1	1.0	0.0	60.5	38.3	0.1	0.9
<b>France Aquitaine</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Biarritz	12	37.2	61.2	1.0	0.5	57.0	41.9	0.5	0.3
<b>Suisse</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Nyon	12	39.0	59.9	0.2	0.7	52.7	46.7	0.1	0.3
<b>France Bourgogne</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Dijon	8	39.3	59.8	0	0.8	61.5	37.2	0.3	0.8
<b>Belgique</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Gembloux	12	41.5	57.9	0	0.5	52.3	47.5	0.0	0
Liège	12	39.4	59.6	0.5	0.5	48.8	51.0	0	0.1
Tournai	12	38.0	61.7	0.2	0	55.5	44.2	0.1	0
<b>Antilles</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Ile de la Réunion	9	40.2	59.4	0.3	0	51.2	48.2	0.5	0
<b>France centre</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Vendée	8	41.6	56.8	0.7	0.75	54.6	44.5	0.1	0.7
<b>Afrique subsahar.</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Abidjan	14	37.9	61.8	0.2	0	55.5	44.4	0	0
Bamako	13	48.0	51.5	0.4	0	50.6	49.3	0	0
Burkina Faso	12	42.9	57.0	0	0	56.3	43.3	0.1	0.1
Cameroun	5	49.3	36.2	13.7	0.62	60.1	37.9	1.1	0.7
RCA Bangui	13	51.7	46.4	1.3	0.51	51.2	48.2	0.2	0.2
Sénégal Dakar	18	42.5	57.4	0	0	45.4	54.5	0	0
<b>Maghreb</b>	NL	NR	RE	NE	autre	NR	RE	NE	autre
Chlef	12	96.4	3.5	0	0	98.7	1.2	0	0

Tableau 1. 7 % de liaisons en lecture et en conversation par régions. Données du projet PFC en ligne.

Mallet (2008 : 190) observe qu'il n'y a pas de différence significative entre la production guidée (45.9%) et la conversation libre (44.1%). C'est sans doute la raison pour laquelle les résultats des deux contextes de production ont été réunis dans la catégorie « conversation ». Par contre, on remarque des différences entre la conversation (comprenant donc la conversation guidée et libre) et la lecture dans nos tableaux 1.7. et 1.8. : les locuteurs réalisent un taux plus élevé de liaisons en lecture.

Dans le tableau 1.8. réalisé à partir des productions d'hommes et de femmes francophones sur le site du projet PFC, les femmes produisent plus de liaisons que les hommes, mais la différence entre lecture et conversation n'apparaît pas.

conversation		lecture	
hommes	femmes	hommes	femmes
44.64	55.35	44.64	55.35

**Tableau 1. 8 % de réalisation de la liaison hommes/femmes France à partir des données du projet PFC en ligne.**

Enfin, il semble que les résultats des productions dépendent de nombreux facteurs, du texte lu, selon qu'il s'agisse de lecture ou de parole spontanée, de l'âge des locuteurs, de leur pays d'origine et de l'évaluation des liaisons et des contextes possibles de la liaison effectuée.

### 3. Les liaisons et le MBU

La fréquence joue un rôle important dans l'évolution de la langue. Les liaisons qui ont survécu à la tendance à ne pas prononcer les consonnes finales sont celles qui se trouvent avec les mots les plus fréquents, les mots au taux de cooccurrence élevé et les mots qui se trouvent dans des structures fréquentes. Ce que Delattre appelle « une union forte » (1947) est ce que nous appelons la fréquence de cooccurrence :

La liaison est la survivance de quelques enchaînements de consonnes finales de l'ancien français. A une époque antérieure, toutes les consonnes finales que l'on écrivait étaient prononcées (ah! le bel âge!). Aujourd'hui, ces consonnes graphiques sont en grande majorité muettes dans les mots isolés; mais dans la chaîne parlée, on les prononce encore quand l'union du mot à consonne finale avec le mot suivant à initiale vocalique a été assez forte pour conserver, à travers les siècles, l'enchaînement ancien. (Delattre, 1947 : 148)



Les combinaisons de mots sont stockées lexicalement et leur fréquence d'utilisation en détermine la force, c'est-à-dire le pourcentage de liaisons réalisées (Bybee, 2002 : 218). C'est la raison pour laquelle les liaisons qui ont été conservées apparaissent dans des contextes syntaxiques les plus fréquents tels que [dét. + nom] et avec les mots les plus fréquents.

Toutefois, lorsqu'on parle de fréquence, on ne parle pas seulement de la fréquence du mot1 et du mot2 mais aussi de la fréquence des structures syntaxiques, de la fréquence des phonèmes de liaison, et du taux de cooccurrence des deux mots. De plus, comme nous l'avons vu dans l'explication du MBU, deux facteurs sont liés aux facteurs de fréquence : la productivité d'une structure (le nombre d'items qu'elle regroupe) et sa complexité (la difficulté à apprendre et maîtriser des règles). C'est la raison pour laquelle il convient d'étudier plusieurs types de fréquence, ainsi que la productivité des structures.

Considérons chaque facteur isolément, puis nous les considérerons ensemble pour créer une analyse multidimensionnelle de la liaison basée sur le MBU.

### **3.1. La fréquence des structures syntaxiques**

Dans le MBU, la fréquence des structures (*type frequency*) est parfois considérée comme plus importante que la fréquence des items (Bybee, 2001), parce qu'une structure regroupe des items, alors qu'un item isolé ne peut représenter que lui-même.

Delattre (1955) dresse une liste de quelques types de structures syntaxiques et évalue leur « union » qu'il classe par ordre décroissant sur une échelle allant de 10 pour la plus forte à 1 pour la plus faible :

- 10: Entre le déterminatif et le nom: des enfants
- 10: Entre l'adjectif et le nom: de beaux enfants
- 10: Entre le pronom personnel et le verbe: ils ont
- 10: Entre le verbe et le pronom personnel: ont-ils
- 9 : Entre l'adverbe et le modifié: tellement aimable
- 8: Entre la préposition et son complément: pendant un jour
- 7: Entre l'auxiliaire et le participe passé: vous avez aidé
- 6: Entre l'auxiliaire et l'infinitif: vous allez aider
- 5: Entre le nom et l'adjectif: des enfants intelligents

- 4: Entre le verbe et son complément: il désirait un cadeau
  - 3: Entre le pronom et le verbe: les miens attendront
  - 2: Entre le nom et le verbe: les enfants attendront
  - 1: Entre la conjonction et ce qui suit: pourtant il était là
  - 1: Entre la conjonction et ce qui précède: il sortait et ne rentrait plus.
- Delattre (1955 : 45)

Delattre est l'un des premiers chercheurs à avoir tenté d'établir des critères autres que syntaxiques, en proposant les premières classifications de liaisons qui ont un rapport avec la fréquence. Ce qu'il appelle le « degré d'union » se base principalement sur le rapport qu'ont les catégories syntaxiques entre elles et aussi le rapport qu'ont les mots entre eux. Plus le degré est fort, plus les mots et catégories syntaxiques ont de chances d'être unis.

En fait, ce « degré d'union » regroupe plusieurs types de fréquences : la fréquence du contexte syntaxique, la fréquence des mots<sup>1</sup> et la fréquence de cooccurrence des mots et aussi, mais dans une moindre mesure, la fréquence des mots<sup>2</sup>. Les articles et les noms sont les catégories syntaxiques les plus fréquentes. Deux catégories syntaxiques très fréquentes ont plus de chances de se trouver ensemble que deux catégories rares (une conjonction + ...).

Cela contribue ainsi à la fréquence de cooccurrence qui détermine la fréquence à laquelle deux mots se trouvent ensemble. Ainsi, un article + nom, représente le contexte syntaxique [art. + nom], qui est le contexte syntaxique le plus fréquent. D'autre part, les articles sont les mots les plus fréquents (mots<sup>1</sup>). De plus, les noms ont le plus de chances de se trouver avec des articles, ce qui veut dire que leur taux de cooccurrence sera également le plus élevé. Enfin, les noms sont plus fréquents que les autres sortes de mots, ce qui veut dire que les mots<sup>2</sup> seront aussi plus fréquents. Autrement dit, le contexte [art. + nom], détermine la fréquence syntaxique, la fréquence de cooccurrence, la fréquence du mot<sup>1</sup> et la fréquence du mot<sup>2</sup>.

En ce qui concerne le rôle de la fréquence syntaxique, Boula de Mareüil, Adda-Decker et Gendner (2003 : 1330-1331) précisent que dans le corpus BREF contenant 66 500 phrases lues par 120 sujets, un peu plus de 25% des 26 000 mots ont des contextes de liaisons

possibles. Leur corpus comporte 91 126 occurrences de contextes de liaisons possibles parmi lesquelles 40 940 liaisons ont été observées, soit un taux global de 45% de liaisons réalisées. Le tableau 1.9. regroupe les sites de liaisons possibles, qui ont été classés par nombre décroissant d'occurrence. Cela nous renseigne sur la fréquence des structures.

Structure syntaxique	occurrences	% données
dét. + nom	14077	15.3
verbe + prép.	6242	6.8
verbe + dét	5797	6.3
verbe + verbe	5225	5.7
nom + adj.	4627	5.0
nom + verbe	4609	5.0
nom + prép.	4266	4.7
nom + conj.	4217	4.6
pron. + verbe	3430	3.7
verbe + adv.	3101	3.4
prép. + nom	2383	2.6
num. + nom	1915	2.1
adv. + adj.	1864	2.0
adv. + verbe	1771	1.9
adj. + num.	1733	1.9

**Tableau 1. 9. Fréquence des structures d'après Boula de Mareüil, Adda-Decker et Gendner (2003 : 1331)**

On remarque par exemple, que le contexte [adj. + nom] est absent de cette liste, mis à part les adjectifs numéraux. En effet, les adjectifs se plaçant toujours avant le nom sont rares, et même ces adjectifs peuvent se trouver dans une autre position (*un homme gros*, plutôt qu'*un gros homme*).

De même, dans le corpus PFC, Durand et Lyche (2008) soulignent que le contexte [adj. + nom] est rare alors que le contexte [nom + adj.] est beaucoup plus fréquent :

Prenominal adjectives appear in large numbers in the base, but only rarely in a liaison environment, and when they do, the liaison is not categorical. The adjective *gros* will serve to illustrate this point: we record 139 occurrences of *gros* in the base, but only 8 in a liaison context, equally distributed between liaison and no liaison. The liaison is then present with the plural form of the adjective (*gros [z]ouvrages*), but not in the singular (*gros// immeuble*). (Durand et Lyche, 2008 : 45)

Durand et Lyche (2008), à partir des données du PFC, confirment les classifications de Delattre concernant les liaisons « obligatoires » (avec dét + nom, etc.), sauf sur trois points.

Les liaisons ne sont pas systématiques après les prépositions monosyllabiques, les adjectifs antéposés et la construction impersonnelle *c'est*.

Nous voyons aussi dans leur classification qu'il est possible de faire des liaisons entre un verbe et un complément, par exemple dans les contextes [verbe + prép.] (Ex : ils arrivent\_en avance) et [verbe + dét.] (Ex : il veut\_un cadeau).

Gadet (1989 : 52-53) établit une liste de certains contextes de liaisons (ils ne sont pas tous présents étant donné que l'auteure se concentre sur la variation, ce qui exclut la plupart des liaisons obligatoires réalisées par tous les locuteurs) qu'elle nomme « les plus communes », « un peu plus rares », « encore plus rares » et « exceptionnelles » en considérant en priorité les facteurs syntaxiques (après une préposition, entre un verbe et un adverbe, etc.), mais aussi les facteurs morphologiques (+ adj. sing.) et phonétiques (après un r de l'infinitif). Elle établit (tableau 1.10.) une hiérarchie des liaisons les plus fréquemment réalisées (entre une préposition et un nom) aux liaisons les plus rarement réalisées, habituellement considérées comme des liaisons interdites [nom + adj. sing.] :

% de LR	exemple	catégorie syntaxique
95 % :	dans un village	entre une préposition (dans) et un nom
95 % :	quand ils	après une conjonction monosyllabique
85 % :	haut en bas	expression figée
60 % :	fait entrer	semi-auxiliaire + verbe
55 % :	mitraillettes allemandes	nom + adjectif au pluriel
40 % :	jamais existé	entre un adverbe et un verbe (participe passé)
30 % :	attendant en silence	entre un participe présent et une préposition
27 % :	s'achever aujourd'hui	après un r de l'infinitif
20% :	réunit aujourd'hui	entre un verbe et un adverbe
12 % :	entrer ici	après un r de l'infinitif
10 % :	enterrer en secret	après un r de l'infinitif
7 % :	accent invisible	nom + adjectif singulier

**Tableau 1. 10. Classement des structures syntaxiques d'après Gadet (1989 : 52-53)**

Si l'on compare ces résultats à ceux de Boula de Mareüil et Adda-Decker et Gendner (2003 :1330) (et accessoirement à ceux de Mallet<sup>20</sup>) dans le tableau 1.11., on voit qu'il y a des tendances similaires. Les liaisons avec les prépositions et conjonctions monosyllabiques sont réalisées de manière quasi systématique (entre 90 et 95% de réalisation). Le contexte nom + adj. sing qui est censé être un contexte de liaison interdite est produit entre 7% et 10%

	Gadet (1989)	BDM, AD et G (2003)	Mallet (2008)
Préposition monosyllabique (dans)	95%	90.5%	91.91%
conjonction monosyllabique (quand)	95%	94.6%	
nom + adjectif singulier	7%	10.4%	
adverbes monosyllabiques			35.99%

**Tableau 1. 11. % de liaisons Gadet (1989), Boula de Mareüil, Adda-Decker et Gendner (2003), Mallet (2008)**

L'observation de corpus de parole adulte (Booij et de Jong, 1987) révèle que les liaisons ne sont réalisées à 100 % que dans quatre contextes : après déterminant, après pronom préverbal, entre verbe et pronom (*prenez-en !*) et dans certaines expressions figées (*tout à l'heure*). Or, ces contextes apparaissent dans les structures les plus fréquentes et ce sont dans ces contextes-là que le taux de cooccurrence des mots est aussi le plus élevé (Bybee, 2001 : 174).

L'étude de Ahmad (1993) souligne également que dans 20 h de parole adulte, les contextes obligatoires sont plus fréquents que les contextes facultatifs. Les liaisons obligatoires appartiennent aux catégories syntaxiques les plus fréquentes : [dét. + nom], [clitiques + verbe] ; les liaisons facultatives appartiennent à des classes plus restreintes et se trouvent donc dans des structures moins fréquentes, comme l'indique Bybee (2005) :

Les liaisons optionnelles font intervenir soit un suffixe de pluriel entre le nom et l'adjectif ou entre l'adjectif et le nom, soit la consonne finale d'un petit nombre d'adjectifs antéposés au nom, soit certaines prépositions, certains adverbes, particules et expressions figées. Dans chacun de ces cas, c'est une catégorie lexicale restreinte qui est concernée. (Bybee, 2005 : 25)

<sup>20</sup> Par « accessoirement » nous entendons que les chiffres sont incomplets et ne sont ici présents qu'à titre indicatif. Le corpus PFC ne réunit pas les conditions syntaxiques. Les chiffres ci-dessus sont ceux que Mallet a pu extraire à partir de certains items.

### 3.2. La fréquence des mots monosyllabiques

La plupart des règles ou études sur les liaisons indiquent qu'elles sont obligatoires après les prépositions, adverbes et conjonctions monosyllabiques, mais qu'elles sont facultatives après les prépositions et adverbes polysyllabiques.

Il convient aussi de préciser que les liaisons sont réalisées plus souvent lorsque le mot1 est monosyllabique, en conversation comme en lecture. C'est en conversation que la différence est la plus visible (les liaisons sont presque trois fois plus souvent produites en conversation si le mot est monosyllabique). Comme le confirme Howard (2005 : 149), il y a une tendance à faire plus de liaisons si le mot1 est plus court que le mot2.

Le tableau 1.12., réalisé à partir des données du projet PFC pour les locuteurs L1, indique que le nombre de syllabes est un facteur majeur dans la réalisation de la liaison.

conversation		lecture	
monosyllabe	plus. syllabes	monosyllabe	plus. syllabes
97.82	2.18	92.49	7.51

Tableau 1. 12. % de réalisation de la liaison par nombre de syllabes (d'après projet PFC)

Certains mots monosyllabiques, tels que les prépositions constituent un facteur dans la production de liaison pour plusieurs raisons. D'une part, les prépositions font corps avec le mot suivant (ce qu'on voit dans des langues à cas où elles peuvent se traduire par un affixe marqueur de cas, en plus des prépositions elles mêmes, comme l'ablatif en latin, ou le prépositionnel en russe). D'autre part, les mots monosyllabiques sont plus simples à retenir et à produire (il est plus rapide de dire « très » que « extrêmement »). Enfin, ils sont plus fréquents. 100% des liaisons du corpus en /n/ se trouvent dans des mots monosyllabiques, 96.56% des liaisons en /z/ et 73.61% des liaisons en /t/ (Mallet, 20008 : 262). Le nombre de syllabes n'influence pas la liaison de la même façon selon les phonèmes de liaison, car plusieurs facteurs (phonémiques et syntaxiques) influent en même temps sur la liaison.

Ainsi, si le mot est monosyllabique et se termine par /n/ (et qu'il s'agit probablement du contexte [dét. + nom] ou [adj. + nom]), la liaison aura beaucoup plus de chances d'être réalisée que si le mot est monosyllabique et se termine par /t/ (et qu'il s'agit probablement du contexte [forme verbale + ...]). Il est ainsi difficile de séparer les facteurs syllabiques, lexicaux, phonémiques et syntaxiques.

### 3.3. La fréquence des phonèmes de liaisons

Après avoir considéré la fréquence des structures syntaxiques et la fréquence des mots monosyllabiques, considérons à présent la fréquence des phonèmes de liaison, pour déterminer son importance dans la réalisation de liaisons.

Dans le projet PFC, Mallet (2008 : 256) distingue :

1. Les contextes possibles, ce qui établit la fréquence potentielle des consonnes de liaison ;
2. Le nombre total et le pourcentage réalisé pour chaque consonne de liaison ;
3. Le pourcentage des liaisons réalisées parmi toutes les liaisons produites.

Les contextes possibles, le nombre de liaisons réalisées et le pourcentage pour chaque consonne ont été classés par ordre décroissant de contextes possibles dans le tableau 1.13.

Consonne de liaison	1. contextes possibles	2. L. réalisées	3. % réalisé	4. % parmi toutes liaisons
/z/	13852	5960	43 %	14.7%
/t/	8113	1188	23.3%	38.7%
/n/	5512	4970	90.2%	46.4%
/R/	942	12	1.3%	0.2%
/p/	98	11	11.2%	
/k/	2	0	0%	

**Tableau 1. 13. Contextes de liaisons possibles, nombre et pourcentage de liaisons réalisées d'après Mallet (2008 : 221, 256)**

Selon que l'on considère le nombre global ou le pourcentage de liaisons réalisées, on n'obtient pas le même classement. Si l'on considère le nombre total de liaisons réalisées, on obtient le classement suivant : /z/ : 5960, /n/ : 4970, /t/ : 1188, /R/ :12, /p/ : 11. Mais si l'on

considère le pourcentage de liaisons réalisées pour chaque catégorie, on obtient le classement suivant : /n/ : 90.2%, /z/ : 43 %, /t/ : 23.3%, /p/ : 11.2%, /R/ :1.3%. Enfin, si l'on considère les pourcentages des liaisons réalisées parmi les liaisons produites, on obtient alors : /n/ : 46.4%, /t/ : 38.7%, /z/ : 14.7%, /p/ et /R/ réunis : 0.2%.

Nous voyons donc que lorsqu'on parle de fréquence des phonèmes de liaisons, il peut s'agir en fait de trois catégories différentes : la fréquence des contextes possibles, le nombre de liaisons réalisées pour chaque phonème et le pourcentage de liaison parmi l'ensemble des liaisons produites. Cela explique aussi en partie les résultats différents selon les études.

Mallet (2008) précise que la fréquence de distribution est la suivante :

Proposition 1 [z] > [n] > [t]	Proposition 2 [z] > [t] > [n]	Proposition 3 [n] > [z] > [t]
Green et Hintze (2001 : 34) Pagliano et Laks (2005 : 5) Durand et Lyche (2008 : 58)	Léon (1992 : 152)	Malécot (1975 : 164)

**Tableau 1. 14. Propositions de distribution des CL /n/, /t/, /z/ Mallet (2008 : 205)**

Dans l'étude de Boula de Mareüil, Adda-Decker et Gendner (2003 :1329), sur un corpus de texte lu, les liaisons les plus réalisées sont par ordre décroissant de fréquence: /z/, /t/, /n/, /R/, /p/. La liaison avec /t/ est plus fréquente qu'avec /n/, en accord avec Léon (1992 :152). Ainsi, la hiérarchie des consonnes de liaison dépend de ce qu'on considère : leur fréquence potentielle, leur fréquence totale (le nombre réalisé pour chaque consonne de liaison), ou la fréquence relative (par rapport aux autres consonnes de liaisons).

### 3.4. La fréquence lexicale

Parmi les facteurs contribuant à la production de liaison, le facteur lexical est sans doute celui qui a été le moins étudié. On sait que plus des séquences sont fréquentes, plus les liaisons sont réalisées (Bybee, 2001), mais pourquoi est-ce le cas ?

Plusieurs facteurs jouent un rôle dans la production de liaisons. Dans l'étude de Malécot (1975), le facteur phonétique est important. On fait plus de liaisons avec certaines



consonnes qu'avec d'autres. En fait, les liaisons se font plus facilement dans des contextes qui regroupent plusieurs paramètres. Par exemple, dans l'étude des prépositions de Malécot (1975), la liaison se fait en plus grand nombre avec

1. les prépositions les plus fréquentes,
2. les prépositions monosyllabiques,
3. les prépositions dont le phonème précédant la consonne de liaison est une nasale (en, dans, sans).

Mallet (2008 : 103-106) arrive aux mêmes conclusions en ce qui concerne le facteur phonétique (nasalité) et le facteur syllabique (la longueur du mot), mais ne tire pas de véritables conclusions sur le rôle de la fréquence, car celui-ci n'a pas vraiment été analysé.

Un exemple de fréquence de cooccurrence est avec *un*. Le taux de cooccurrence est lié à la structure syntaxique. S'il est employé comme article (c'est donc une structure plus fréquente) la liaison se fait dans tous les cas : *un ami, un autre*. S'il est employé comme pronom, dans une expression (presque figée) : *L'un à l'autre, l'un avec l'autre* (Morin et Kaye, 1982) la liaison est facultative (le taux de cooccurrence est plus faible). Dans tous les autres cas (où *un* est un pronom), la liaison est impossible car les structures sont plus rares : *ils sont venus, l'un avec sa mère, l'autre avec son père* (Bybee, 2005 : 29). On voit donc que ce n'est pas seulement la fréquence du mot (ou du mot 2), mais aussi la fréquence de la structure syntaxique qui joue un rôle dans la production de liaison.

Ågren (1973) remarque que les liaisons sont obligatoires dans certaines expressions figées alors qu'elles ne le sont pas dans des expressions moins fréquentes. Durand et Lyche (2008 : 57) remarquent que dans les séquences *les Nations-Unies, les Jeux Olympiques*, la liaison est obligatoire, alors que dans les autres contextes [nom plur. + adj.] la liaison apparaît dans 26% des cas. Notons au passage que cette même expression figée *les Jeux Olympiques*, censée être obligatoire, n'est pas réalisée par 14 des 195 locuteurs de l'enquête du projet PFC.

Delattre (1947 :156-157) propose une liste assez longue d'expressions figées où la liaison est obligatoire (*un pied à terre, ses faits et gestes, monts et merveilles, tout au plus, mot à mot, du tout au tout*, etc.). Ces expressions sont elles-mêmes plus fréquemment utilisées que d'autres expressions figées, qui elles n'ont pas de liaisons obligatoires.

Selon Bybee (2001 :184-185), c'est la fréquence d'un mot (*token frequency*) d'une séquence particulière qui est la plus importante pour résister à la régularisation. Ce qu'on appelle cohésion syntaxique, c'est la fréquence de cooccurrence des mots, et c'est ce qui détermine la force de l'association entre le premier élément et le second. Ces connections sont stockées en mémoire et renforcées par leur usage fréquent.

Selon Bybee (2005, 24-25), la liaison est le résultat direct de la fréquence de cooccurrence : les mots souvent employés ensemble semblent liés par une plus forte cohésion. Des mots et des suites de mots sont mémorisés dans le lexique. Le critère de mémorisation est la fréquence d'usage et ceux qui sont employés fréquemment voient leur force lexicale augmenter, sont plus facilement accessibles que les autres et résistent mieux aux changements par analogie.

Ce que Bybee (2001) appelle « le facteur de cooccurrence » est similaire à ce que Delattre (1955) appelle : « le degré d'étroitesse » entre les mots. On peut dire, en accord avec Delattre (1947, 1955, 1956), Bybee et Scheibman (1999) et Bybee (2001, 2002, 2005) que la fréquence de cooccurrence des mots et la fréquence des mots joue un rôle majeur dans la production de liaisons. Plusieurs études appuient cette hypothèse.

Adda-Decker, Boula de Mareüil et Lamel (1999) ont aussi abordé l'effet de la fréquence sur les liaisons. Dans leur corpus de 155 heures de parole, ils ont trouvé que, si l'on considère les 256 mots les plus fréquents, 75% des liaisons sont réalisées, alors que si l'on étend l'échantillon aux 2560 mots les plus fréquents, le taux de réalisation diminue à 64%, contre 55% si l'on considère tout le corpus.

En calculant l'entropie<sup>21</sup>, avec des données quantitatives sur la fréquence des mots les uns par rapport aux autres, Culetto (2007) met en évidence le fait que la fréquence joue un rôle plus important et plus fiable que la syntaxe :

The results of the correlation test confirmed the hypothesis that liaison is, on the whole, linked to conditional entropy, even if many individual lexical items do not conform to this prediction. As it is generally considered that liaison occurs between words that are strongly connected, my research would imply that this cohesion can be measured by conditional entropy. From the point of view of formal language theory, a simple probabilistic finite-state machine is sufficient to account for this phonological phenomenon, without the need of more elaborate syntax. (Culetto, 2007: 65)

Boula de Mareüil et Adda-Decker (2002 : 2276) remarquent que la réalisation des liaisons dépend des mots ; c'est la raison pour laquelle ils se réfèrent à la catégorie « entrée lexicale ». Ils constatent qu'environ la moitié des liaisons possibles sont réalisées dans BREF (55.3%) et un peu moins dans MASK (42.8%). Ils ont trouvé que la liaison est plus fréquente avec les mots les plus fréquents (75%) qu'avec les mots rares (20%), dans le corpus BREF (lecture de journaux). Dans le corpus MASK (parole guidée), les 10 mots les plus fréquents couvrent 30% du corpus et 20% du corpus de BREF. Les 100 mots les plus fréquents couvrent 80% de MASK et moins de 50% de BREF. Cela indique que la fréquence des mots se remarque davantage dans la parole spontanée que dans la lecture, car il y a plus de répétitions en parole spontanée, les mots répétés sont les mots les plus fréquents car il y a moins de recherche stylistique.

Nous voyons donc que la réalisation des liaisons met en œuvre divers paramètres qui sont tous interreliés : les « entrées lexicales » dépendent des conditions d'énonciation, du thème discuté, de la longueur des mots, de la fréquence des mots, et des locuteurs eux-mêmes. Ceci est d'autant plus le cas pour les liaisons facultatives.

Le nombre de liaisons réalisées, et surtout le taux de liaisons facultatives réalisées est en partie lié aux compétences des locuteurs en lecture et en écriture (*literacy*). Certes, les

---

<sup>21</sup> Entropie : Dans la théorie de la communication, nombre qui mesure l'incertitude de la nature d'un message donné à partir de celui qui le précède. (L'entropie est nulle quand il n'existe pas d'incertitude.) Définition du Petit Robert électronique 2007.

liaisons les plus fréquentes (les liaisons obligatoires) sont apprises avant que l'enfant ne sache lire et écrire. L'enfant reproduit ce qu'il a entendu dans son input et pour les liaisons les plus fréquentes de type [dét. + nom], les connaissances du système graphique ne sont pas nécessaires.

Par contre, pour les liaisons facultatives moins fréquentes, l'une des conditions pour (bien) les réaliser c'est aussi de connaître l'orthographe et la grammaire. Si l'on ne sait pas comment les mots s'écrivent, on ne peut pas savoir quelle consonne utiliser pour la liaison et surtout s'il faut la faire ou non. Par exemple, pour faire la liaison correctement avec *vous viendrez z-avec moi* ou ne pas faire la liaison avec *ira / à moi* (exemples donnés dans le sketch des Nuls, *les Liaisons Dangereuses*), il faut savoir comment *viendrez* et *ira*, s'écrivent.

C'est la raison pour laquelle un jeune enfant qui ne sait pas encore très bien écrire peut faire des erreurs de liaisons, de type « pas-t-encore » comme on le voit dans les exemples de Pallaud et Savelli (2001 :124-127), et un adulte illettré ou analphabète ferait le même type d'erreurs comme le fait le francophone de Louisiane analphabète en français cité dans Morin (2005).

Bien que les études de Durand et Lyche (2008) et l'étude de Mallet (2008) portant sur le corpus PFC n'aient pas montré de différences selon le niveau d'éducation des sujets, d'autres études mettent en évidence le rôle du niveau socioculturel des locuteurs ou de leur famille (Encrevé 1988, pour les adultes, et Chabanal, 2007 pour les enfants). C'est entre autres parce que les facteurs sociolinguistiques diffèrent largement d'un sujet à l'autre, que les résultats diffèrent d'un sujet à l'autre et d'un groupe à un autre, comme on le voit dans le tableau 1.15. Par exemple, les liaisons facultatives sont réalisées à 34% dans l'étude d'Ashby (1981) contre 61.6% dans l'étude d'Ågren (1973).

	liaisons obligatoires et facultatives réunies			L facultatives	L facultatives
	PFC, conv.	PFC, lect.	Malécot (1975)	Ågren (1973)	Ashby (1981)
% Liaisons enchaînées	43,69	59,51	64	61,6	34

**Tableau 1. 15. % de réalisation de liaisons facultatives d'après différentes études**

Si l'on considère les facteurs sociolinguistiques déterminant la production de liaisons facultatives tels que l'âge, le niveau d'instruction et le nombre d'années d'études, la profession et le statut social du locuteur et de sa famille, ou le lieu d'étude et le type de formation (Encrevé 1988), ces facteurs sont eux-mêmes liés au nombre de liaisons entendues, donc au facteur de fréquence.

Plus une personne est éduquée, plus sa famille est éduquée, plus elle aura entendu de contextes variés et de liaisons et réciproquement. Moins une personne a un statut social élevé, moins elle sera en contact avec des personnes de statut social élevé, moins les conversations en style recherché seront fréquentes, moins elle entendra de liaisons (facultatives) et moins elle sera susceptible d'en produire.

Il en va de même pour le facteur géographique : plus une personne est en contact avec un groupe linguistique qui produit un grand nombre de liaisons, plus cette personne entendra de liaisons et plus elle sera susceptible d'en produire. Par exemple, l'étude du corpus PFC souligne que les Français ont tendance à faire plus de liaisons que les Suisses et les Belges, qui en font eux-mêmes plus que les Africains.

Puisque les enfants imitent la langue de leurs parents, de leurs enseignants et de leur entourage, il paraît logique que, si les Africains entendent moins le français que des Français, ils produisent moins de liaisons qu'eux. En effet, leur langue maternelle n'est pas le français, c'est seulement la langue d'éducation ; d'autre part, si la langue maternelle de ceux qui servent de modèle d'apprentissage du français (les enseignants par exemple) n'est pas le français, cela suppose que ces personnes produiront moins de liaisons que les locuteurs natifs.

A leur tour, les apprenants du français L2 devraient produire moins de liaisons que les locuteurs natifs. C'est en effet le cas dans le projet PFC, puisque les Africains en produisent moins que les Français. Cela a des implications sur les sujets se trouvant en contexte restreint d'utilisation de la langue.

D'après le MBU, plus des structures sont présentes dans la conversation de francophones, plus ils devraient les maîtriser et donc les utiliser. Green et Hintze (1990) concluent que la resyllabation (*linkage*), c'est-à-dire l'enchaînement, a lieu toutes 1.66 secondes et la liaison une fois toutes les 7 secondes. Autrement dit, les enchaînements sont plus de quatre fois plus fréquents que les liaisons. Ils sont dits « obligatoires », alors que ce n'est pas le cas pour toutes les liaisons. Les enchaînements devraient donc être réalisés en grande proportion par les francophones.

Morin (2005) remarque que certains adjectifs peuvent se placer avant le nom, seulement au féminin ou au pluriel : *une chaude saison, une longue histoire, de longs efforts* ou devant une consonne *un grossier personnage*. Par contre, une forme analogue au masculin singulier commençant par une voyelle paraîtrait peu authentique et la liaison serait certainement non produite *un chaud été, un long effort, un grossier r-homme*. Morin (2005) explique qu'on dirait plutôt en parole spontanée, *un été chaud, un grand effort, un homme grossier* pour éviter une liaison que des locuteurs natifs jugeraient impossible, ou tout au moins peu authentiques.

Comment peut-on donc expliquer qu'on dise souvent *un grossier personnage* alors qu'on ne peut pas dire *un grossier homme* ? Là aussi, le MBU fournit une réponse possible qui explique pourquoi une séquence a plus de chance d'être produite et une liaison a plus de chances d'être réalisée, comme dans *grand honneur* par rapport à *grand émoi*, séquences tirées de Durand et Lyche (2008 : 44).

Selon le MBU, le locuteur traite la langue selon une série de probabilités à partir d'une banque de données établie d'après son expérience. Nous pouvons tenter de reproduire une simulation du traitement des données et des probabilités de façon similaire au cerveau humain, à partir des facteurs de fréquence et de productivité que nous avons décrits précédemment. Ces facteurs sont cumulés pour arriver à un total qui peut en partie prédire la

production de liaisons. C'est la raison pour laquelle nous avons étudié plusieurs facteurs de fréquence : fréquence du mot1, du mot2, fréquence de cooccurrence, fréquence de structures syntaxiques, etc.

Pour illustrer ce que nous entendons par un calcul statistique de données à partir de plusieurs facteurs, imaginons un éventail de points allant de 0 à 3 pour chaque facteur concerné. 0 = très rare, 1 = rare, 2 = moyen, 3 = fréquent et très fréquent.

En ajoutant divers facteurs, nous pouvons obtenir un certain nombre de points qui constituerait un indice global de probabilité. Un tel système de codification demande à être raffiné, d'après les facteurs qui auraient plus d'importance que d'autres et représente un système beaucoup plus simple que le fonctionnement réel du cerveau humain. Toutefois, cette hypothèse servira de base à notre réflexion pour tester l'hypothèse du MBU et pour répondre aux questions concernant les liaisons.

Si l'on considère par exemple les deux séquences : *grand émoi* (tableau 1.16.), et *grand honneur* et (tableau 1.17.), séquences mentionnées dans Durand et Lyche (2008 : 44), et que nous analysons la fréquence du mot1, du mot2 et la fréquence de cooccurrence de ces séquences (d'après *Lexique*), ainsi que la productivité des structures phonétiques et syntaxiques (d'après notre analyse de la productivité des structures à partir des explications des études sur le MBU dont nous avons précédemment discuté), nous comprenons pourquoi dans un même contexte syntaxique, la liaison a plus de chances d'être réalisée dans *grand honneur* que dans *grand émoi*.

caractéristiques structure/item	exemple	fréquence	/3
fréquence de la structure	Adj. qual. + nom	rare	1
fréquence du mot 1	grand	très fréquent	3
fréquence du mot 2	émoi	rare (hors liste)	0
complexité de la structure phonétique	complexe (d devient t)	complexe	0
complexité (grapho-phon) du mot 2	simple	fréquent (é)	3
nombre d'items dans la structure syntaxique	classe réduite	seuls quelques adj	0
(productivité de la structure)	peu productive	gros, grand, petit...	
taux de cooccurrence des 2 mots	très faible	0 (sur <i>Lexique</i> )	0
		<b>total</b>	<b>7</b>

**Tableau 1. 16. Fréquence et productivité de la séquence un grand émoi (Durand et Lyche, 2008 : 44)**

caractéristiques structure/item	exemple	fréquence	/3
fréquence de la structure	adj qual + nom	rare	1
fréquence du mot 1	grand	très fréquent	3
fréquence du mot 2	honneur	très fréquent	3
complexité de la structure phonétique	complexe (d devient t)	complexe	0
complexité (grapho-phon) de l'item	complexe	h aspiré vs h muet	0
nombre d'items dans la structure syntaxique (productivité de la structure)	classe réduite peu productive	seuls quelques adj. gros, grand, petit...	0
taux de cooccurrence des 2 mots	assez fréquent	69 (sur <i>Lexique</i> )	2
		<b>total</b>	<b>9</b>

**Tableau 1. 17. Fréquence et productivité de la séquence *un grand honneur* (Durand et Lyche, 2008 : 44)**

Il apparaît d'après ces données de fréquence et d'après notre analyse de la productivité des structures, que la liaison dans *grand honneur* a plus de chances d'être produite que dans la séquence *grand émoi*, mais que le total des deux séquences est relativement bas, ce qui devrait créer de la variation dépendant d'autres facteurs spécifiques au locuteur. En ce sens, nous rejoignons Durand et Lyche (2008 : 1), selon lesquels il faut considérer la variation (sociolinguistique), ce qui dépend de facteurs interpersonnels, en plus d'autres facteurs : « Dealing with liaison requires interaction with all the components of the grammar, while tackling at the same time the quicksands of variation. »

Des facteurs tels que l'âge et le niveau d'étude (d'éducation) du locuteur influencent la production de liaisons. Le milieu socioéconomique et l'éducation des parents jouent un rôle dans l'apprentissage de l'enfant. L'expérience langagière de l'enfant semble conservée à l'âge adulte puisque les enfants issus d'un milieu socioéconomique défavorisé font moins de liaisons que leurs pairs à tous les âges (Dugua, 2005) et les adolescents et les adultes peu éduqués font moins de liaisons que les adultes plus éduqués (Encrevé, 1988). Les locuteurs qui ont des parents francophones et parlent uniquement le français, sont plus exposés à la langue que ceux dont le français est seulement une langue d'éducation. En outre, nous pouvons arriver à un système similaire pour les conditions sociolinguistiques du locuteur.

Par exemple, si l'on imagine deux locuteurs vivant dans un pays où le français est la seule langue (la France), ayant un statut socioéconomique opposé (tableaux 1.18. et 1.19.), et des caractéristiques sociales différentes, nous obtenons un nombre de points différent.



Parmi les critères qui participent à l'input reçu, les plus importants sont ceux en rapport avec la quantité d'input reçu. Mais si les parents parlent une autre langue à la maison, alors l'input sera forcément restreint. De plus, le niveau d'éducation des parents joue un rôle sur la qualité de l'input reçu.

Imaginons que le locuteur X est médecin, a ses deux parents eux-mêmes médecins qui ont toujours parlé le français à la maison dans une langue très châtiée. Il a eu une exposition maximale à la langue. On obtient pour ce locuteur le maximum de points : 9 points.

Par contre, les conditions d'apprentissage et l'exposition à la langue sont différentes pour le locuteur imaginaire Y. Celui-ci a 16 ans, il est issu de milieu défavorisé. Sa mère ne travaille pas et son père est maçon. Ses parents ne parlent le français à la maison que s'ils ont des invités qui ne sont pas de leur communauté linguistique, c'est-à-dire très rarement. Ils ont toujours parlé une autre langue que le français à leurs enfants, avec leur famille et leurs amis. Le locuteur a arrêté l'école avant le bac (son diplôme d'études secondaires), n'a dans son entourage que des personnes au chômage se trouvant dans la même situation linguistique et sociale. Il a donc eu une exposition minimale à la langue à la maison et dans son entourage immédiat. On obtient pour ce locuteur 1 point.

caractéristiques d'un locuteur X	exemple	points/catégorie	
français parlé à la maison	toujours	3	3
	souvent	2	
	parfois	1	
	jamais	0	
niveau d'éducation des parents	bac + 4 et plus	3	3
	bac à bac + 2	2	
	moins que le bac	1	
	analphabètes	0	
statut socio économique du locuteur	très élevé	3	3
	élevé	2	
	moyen	1	
	bas/très bas	0	
	<b>total</b>		9

**Tableau 1. 18. Exemple 1 de caractéristiques sociolinguistiques d'un locuteur X (milieu socioéconomique très élevé) et du lien de ces caractéristiques avec la production de liaisons selon un système de points**

caractéristiques d'un locuteur Y	exemple	points/catégorie	
français parlé à la maison	toujours	3	
	souvent	2	
	parfois	1	1
	jamais	0	
niveau d'éducation des parents	bac + 4 et plus	3	
	bac à bac + 2	2	
	moins que le bac	1	
	analphabètes	0	0
statut socio économique du locuteur	très élevé	3	
	élevé	2	
	moyen	1	
	bas/très bas	0	0
	<b>total</b>		1

**Tableau 1. 19. Exemple 2 de caractéristiques sociolinguistiques d'un locuteur Y (milieu socioéconomique très bas) et du lien de ces caractéristiques avec la production de liaisons selon un système de points**

Ainsi, pour des locuteurs vivant dans un même pays et ayant tous deux le français comme langue majoritaire, les conditions sociolinguistiques et les conditions d'acquisition de la langue varient grandement. Les critères personnels pourraient aller de 0 points à 9 points.

Si l'on cumule les caractéristiques sociales d'un locuteur aux caractéristiques d'une séquence de mots (tableaux 1.20. et 1.21.), on arrive à des résultats différents, puisque les conditions du locuteur varient et les séquences de mots varient également.

	caract locuteur	séquence grand honneur	total
caract. X	9	9	18
caract. Y	1	9	10

**Tableau 1. 20. Exemple – cumul fréquence et productivité *grand honneur* + critères personnels**

	caract locuteur	séquence grand émoi	total
caract. X	9	7	16
caract. Y	1	7	8

**Tableau 1. 21. Exemple – cumul fréquence et productivité *grand émoi* + critères personnels**

On peut imaginer un seuil minimum au dessus duquel une liaison est produite, par exemple 10 points. Si le nombre de points est supérieur à 11, la liaison est produite, si le nombre de points est entre 9 et 10, la liaison est produite mais avec une erreur ou une hésitation ou reprise; et enfin, si le nombre de points est inférieur à 9, elle n'est pas produite.

Alors le locuteur X produira les liaisons dans *grand honneur* et *grand émoi*, alors que le locuteur Y en produira dans *grand honneur* avec une erreur, une hésitation ou une reprise, mais pas avec *grand émoi*. La liaison a plus de chances d'être produite avec *grand honneur* qu'avec *grand émoi*, mais le contexte *grand honneur* peut tout de même provoquer des erreurs (de segmentation, mauvaise consonne de liaison, etc.) et des non-liaisons chez certains locuteurs et d'autant plus ceux pour qui le français n'est pas la langue maternelle). Nous souhaitons avec ces deux exemples réfléchir aux causes de la variation sociolinguistique et tentons d'expliquer pourquoi certaines séquences sont réalisées par certains locuteurs avec une liaison, d'autres avec des hésitations ou des erreurs et d'autres ne sont pas réalisées.

Avec ces deux exemples, 6 locuteurs n'ont pas fait la liaison avec *grand* dont 4 cas avec *grand émoi* et deux personnes ont prononcé [d] au lieu du [t] attendu (Durand et Lyche, 2008 : 44) : « For *grand émoi*, *grand honneur*, we assumed that in a reading task, all 100 speakers would link the adjective to the following noun, which is not what we observe: six speakers do not make the liaison and two pronounce a [d] instead of the expected [t]. Four of the six cases concern *grand émoi*, suggesting that the lack of familiarity with the construction impacts on the absence of liaison. »

Le facteur lexical (couplé aux facteurs sociaux) permet de raffiner l'analyse et de prédire avec plus de précision les liaisons qui peuvent être produites, et ce d'autant plus si l'on connaît le bagage linguistique du locuteur. En fait, on peut prédire les productions de liaisons, d'après l'input reçu par les locuteurs, ce qui comprend les différents types de fréquences, ainsi que les conditions d'apprentissage et l'exposition à la langue du locuteur.

Nous en venons à la même conclusion que Durand et Lyche :

Future analyses will surely benefit from abstract model construction and a better understanding of the connections between the various 'levels' of description; but they will have to take explicitly into account the results of sociolinguistic surveys, acquisition studies, experimental phonetics as well as psycho-(neuro-) linguistic investigations, including the relationship between speech and writing. As stressed in Chevrot, Fayol and Laks (2005), these will have to acknowledge that French liaison is not a homogeneous locus but a multifaceted phenomenon requiring us to accept, without demur, the crossing of disciplinary boundaries. Above all, future accounts will have to accept that intuition and recourse to past authorities are not sufficient to build a solid picture of French liaison. (Durand et Lyche, 2008 : 61)

#### 4. Les liaisons « hors-normes » et le MBU

Selon les définitions traditionnelles des liaisons, une liaison s'effectue si le mot1 se termine par une consonne latente qui est prononcée et enchaînée au mot2 suivant, si celui-ci commence par une voyelle ou un h aspiré.

Les définitions et catégories de liaisons n'étant pas clairement établies, nous pouvons rajouter à la définition ci-dessus que la liaison se fait dans certaines conditions syntaxiques, phonétiques et lexicales qui correspondent à la norme reconnue et utilisée par la majorité des locuteurs du français standard. D'après une telle définition, les personnes parlant une variété du français autre que celle de la majorité des locuteurs du français standard, vont donc utiliser une langue « hors-norme ». Il nous reste donc à déterminer ces productions « hors-normes ».

Étant donné la complexité du système de liaison et du système de syllabation ouverte, les francophones savent très jeunes qu'il est préférable d'éviter l'hiatus. Nous avons vu au chapitre précédent qu'au début, les enfants ne réalisent pas toutes les liaisons obligatoires et font beaucoup d'erreurs (Dugua, 2005, 2006). A partir de 6 ans, ils réalisent toutes les liaisons obligatoires, mais font encore peu de liaisons facultatives et continuent à faire des erreurs (cuirs et velours) avec celles-ci. Ils réalisent des liaisons interdites jusqu'à un âge avancé. Or, toutes ces erreurs sont directement liées à la fréquence des structures (grammaticales ou phonétiques), à la fréquence du mot2 et à la fréquence de cooccurrence du mot1 et du mot2.

En outre, l'un des facteurs qui participe grandement aux erreurs de liaisons est la contrainte anti-hiatus. Dès l'âge de 6 ans, les enfants tentent d'éviter l'hiatus en rajoutant une consonne épenthétique, même s'ils n'ont jamais entendu de telles séquences, comme l'illustrent les exemples de Pallaud et Savelli (2001 : 124) : « Souris qui-z-habitent (6 ans) [...] des petits qui-z-embêtent les autres (6 ans) [...] ceux qui-z-ont pas cantine (6 ans) [...] »

Pallaud et Savelli (2001 :126) ont comparé les erreurs d'enfants à celles d'adultes pourtant spécialistes de la parole tels que des journalistes et des présentateurs d'émissions

télévisées ou radiophoniques. Ils ont remarqué des erreurs similaires chez les enfants et chez les adultes, et en concluent qu' : « on pourrait en faire une collection à tous les âges ».

Comme en témoignent les exemples de Pallaud et Savelli (2001 : 126), les erreurs de liaisons des adultes consistent soit à rajouter une consonne de liaison, soit à prononcer une consonne qui ne correspond pas à ce qu'on attend dans cette situation, notamment à cause de la consonne écrite : « Constitue-t-un point... (homme politique, 2000) [...] jeudi-z-après midi (homme politique, 2000) [...] Vingt et un-z-otages (radio, 2000) [...] Averses parfois-t-orageuses (météo, radio) (journalistes, radio, 2000) [...] »

Toutefois, les erreurs de liaisons ne sont pas aléatoires ; elles sont motivées par des règles, par des principes sous-jacents. Aucune étude n'a fait état de liaisons obligatoires non effectuées chez des adultes ou chez des enfants de plus de 6-7 ans, si ce n'est des erreurs de parcours lors d'hésitations et reprises. Aucune étude n'a fait état de liaisons fréquentes non réalisées ou mal réalisées par des enfants de plus de 8 ans. Les liaisons fautives (liaison avec la mauvaise consonne) sont toujours réalisées avec des liaisons facultatives moins fréquentes ou avec des structures complexes ou particulières, ou avec des structures comprenant une classe restreinte d'items telle que les structures avec h aspiré.

Les erreurs de liaisons des enfants comme des adultes ont la même cause et se rencontrent dans les mêmes situations : ce sont des structures et des séquences moins fréquentes, séquences qui n'ont pas été enregistrées dans le lexique des locuteurs. Après quelques définitions, nous verrons des exemples d'erreurs fréquentes et nous les expliquerons à la lumière du MBU.

#### 4.1. Liaisons « hors-normes » et définitions

Les erreurs de liaisons sont commises et reconnues depuis longtemps et suscitent chez les interlocuteurs de vives réactions qui se situent entre le rire, l'énervement et l'indignation, selon l'auteur de l'erreur, le type d'erreur et les circonstances dans lesquelles les erreurs ont été commises. Les exemples d'erreurs de liaisons sont présents dans de nombreux articles, car ces erreurs sont représentatives de plusieurs questions importantes. La première constatation qui s'impose est que les erreurs de liaison semblent être motivées en grande partie par la volonté d'éviter l'hiatus, ce que mentionne Delattre (1955 : 46).

Le terme *hiatus* est défini par le *Dictionnaire de l'Académie Française* (1831 : 694) comme étant un : « Terme emprunté du latin, pour exprimer une sorte de bâillement, qui fait quelquefois un mauvais effet dans la prose, et qui est interdit dans la poésie. L'hiatus est proprement entre une voyelle qui finit un mot, et celle qui en commence un autre, et où il ne se fait point d'élision. Dans les vers, il faut éviter les hiatus. Un hiatus qui blesse horriblement l'oreille. [...] » Si l'hiatus « blesse l'oreille » des francophones, ils vont donc l'éviter. En outre, selon Delattre (1955 : 47), tous les hiatus n'auraient pas le même statut. Les hiatus seraient moins « acceptables » avec deux voyelles de même timbre.

La question de l'hiatus n'est certes pas réservée à la poésie ou à la récitation de vers. On peut noter le même phénomène avec l'apostrophe, les modifications morpho-phonétiques que l'on trouve dans les adjectifs antéposés telles que *cet, bel, fol, mol*, ainsi que celles des adjectifs possessifs et démonstratifs du type *son amie, cet homme*.

On trouve en français de nombreux phonèmes éphelestiques<sup>22</sup> grammaticalement corrects tels que les /z/ ou /t/ euphoniques du type *va-t'en, va-t-il, reste-t-il, vas-y, donnes-en*

---

<sup>22</sup> Un phonème éphelestique (du grec ἐφελκυστικός, « attiré à la suite ») est, en phonétique, un phonème — la plupart du temps une consonne — ajouté à la fin d'un morphème ou d'un mot (c'est dans ce cas une paragoge) pour des raisons euphoniques, lesquelles sont le plus souvent la résolution de l'hiatus. On parle, plus couramment, de « son de liaison » bien que le terme de liaison appliqué au français renvoie à une notion différente : c'est un phonème éphelestique dont l'origine est étymologique (c'est un ancien son qui réapparaît entre deux voyelles). Les phonèmes éphelestiques sont des cas de sandhi.

qui ont pour but d'éviter l'hiatus ou encore l'ajout du *l'* à *on* quand celui-ci se trouve après *et* ou *mais* (*et l'on dit, mais l'on croit*).

A l'oral familier, l'élision permet le phénomène inverse régi par le même principe : éviter l'hiatus. Toutefois, il s'agit ici de simplifier la parole en enlevant une voyelle « gênante ». C'est le cas par exemple avec *tu as* qui devient *t'as*, *qui est* qui devient *qu'est*, *puis* qui devient *pis*, *bien* qui devient *ben*, *lui* qui devient *y* (*on y a mis la tête contre une brique* : dans la chanson de Renaud *Deuxième Génération*) pour n'en citer que quelques-uns.

Alors qu'en français standard on dit *donne-m'en*, en français familier on peut dire : *donne-moi-z-en* (basé sur *donnes-en*). La version familière correspond à « la loi du moindre effort » dans le sens où il n'y a qu'à rajouter une consonne épenthétique devant les pronoms alors que dans la version standard, *moi* devient *m'*, ce qui est unique à quelques structures particulières [moi + en], regroupant une classe restreinte de formes impératives et donc une structure moins productive.

Pour éviter l'hiatus, il y a donc deux solutions : soit on enlève un phonème, soit on en rajoute un. Le problème est que certaines structures sont acceptées et même recommandées avec certains items, alors qu'elles sont considérées comme agrammaticales avec d'autres items. Par exemple, il est grammatical de rajouter un *s* dans les cas suivants : *vas-y*, *cherches-en*, *donnes-en*, alors qu'il est agrammatical de dire *donne-moi-z-en*. Il est grammatical de rajouter un *t* dans les cas suivants : *va-t'en*, *a-t-il*, *cherche-t-il*, alors qu'il est agrammatical de dire *constitue t-un point*. Autrement dit, les règles dépendent de l'usage et de la norme d'un groupe linguistique. Mais comment sait-on si une structure est grammaticalement correcte ?

Si un locuteur a entendu dans son entourage des erreurs de liaisons, qui sont elles-mêmes similaires à d'autres structures pourtant grammaticales, il peut considérer ces structures comme acceptables, car elles répondent aux mêmes exigences (anti-hiatus, ajout d'une consonne épenthétique *t* ou *z*) et sont présentes dans son lexique au même titre que

celles qui sont grammaticales. La question de la grammaticalité dépend donc en partie de ce qui est jugé acceptable, au moins à l'oral, par les personnes de pouvoir qui décident des règles phonétiques, grammaticales et orthographiques, mais aussi par un groupe linguistique.

Ce qui est jugé comme acceptable (sinon correct) dépend de plusieurs facteurs. Par exemple, Morin et Kaye (1982 : 317) rappellent qu'il était avant possible (bien que rare) de faire des liaisons avec /t/, /p/ et /k/ dans le contexte [nom + adj.], alors que ce n'est plus possible de nos jours, dans les cas suivants : *Un soldat-t-anglais, du drap-p-anglais, un sang-k-impur*.

L'acceptabilité de la liaison dépend du « capital linguistique » (selon les termes de Laks, 2000 : 2) de l'interlocuteur, c'est-à-dire de son niveau d'éducation, de son âge, de sa classe sociale, etc. et de conditions géographiques telles que le pays et la région de l'interlocuteur. On trouve par exemple des variantes canadiennes dans lesquelles la liaison avec *ils* avant un verbe à initiale vocalique peut être réalisée sans le [z] du pluriel et prononcé [ijõ] (Durand et Lyche, 2008 : 53).

Selon les termes de Malécot (1975), certaines liaisons sont plus « prestigieuses » que d'autres. De même, Morin et Kaye (1982 : 297) expliquent que certaines erreurs correspondent à différents registres de langue. Certaines erreurs se font plutôt dans la langue familière : *Quatre-z-enfants, donne-moi-z-en, un sacré-z-ennui, un gros-t-enfant* et d'autres dans une langue plus soutenue, telles que le *z* du pluriel après les noms pluriels et le *t* après les verbes.

Selon Morin et Kaye (1982), les erreurs de liaison sont fréquentes dans le discours des personnes moins éduquées, mais aussi dans le discours de certaines personnes éduquées qui dans une situation d'hypercorrection utilisent un registre de langue élevé qu'elles ne maîtrisent pas toujours. Le type d'erreur est différent selon les classes sociales. De même, certaines erreurs de liaisons sont moins acceptables que d'autres et certaines erreurs



représentent un marqueur sociolinguistique d'une classe sociale plus basse. Par conséquent, certaines erreurs sont particulièrement « remarquables », alors que d'autres passent pratiquement inaperçues ce que soulignent Morin et Kaye (1982) :

(Les erreurs de liaison) correspondent à des généralisations des règles des liaisons mais ces erreurs sont absentes du discours de ceux qui savent quand l'orthographe les interdit. [...]  
Parmi les erreurs que l'on trouve typiquement dans le registre soutenu, [...] le *z* après des noms pluriels tels que les chemins de fer-*z*-américains (velours sur le nom) et le *t* observé après les verbes comme dans il devra t-y aller (pataquès). Toutes les erreurs ne sont pas marquées socialement de la même façon ; les velours sur les clitiques et les adjectifs numéraux ne se remarquent pratiquement pas tandis que les cuirs et velours sur les adjectifs sont très fortement marqués socialement<sup>23</sup>. (Morin et Kaye, 1982 : 297)

Quoi qu'il en soit, si l'on considère les erreurs de liaisons d'après les normes des manuels de phonétique ou d'après les règles classiques mentionnées précédemment, il apparaît que celles-ci sont fréquentes, c'est pourquoi l'Académie Française<sup>24</sup> en fait une mise en garde en précisant les différents types d'erreurs à éviter (cuirs, velours et pataquès).

De même, Wikipedia met en garde les locuteurs contre les effets des erreurs de liaison<sup>25</sup> : « [...] pratiquer une liaison là où elle est impossible peut aussi passer pour une erreur [...] : la liaison passera pour l'indice d'un locuteur peu cultivé devant un interlocuteur la pratiquant dans la langue familière, la liaison devant le h « aspiré » (des-*z*-haricots) peut sembler fruste ou humoristique [...]. »<sup>26</sup>

Les fautes de liaisons ne sont pas récentes, pas plus que leurs définitions. La précision dans le type de faute indique que c'est un phénomène important pour les francophones. Le terme *cuir* date de 1783, le terme *pataquès* de 1784 et le terme *velours* date de 1822. À cela, on peut rajouter un terme moins connu : *la psilose*. Le premier sens du mot *psilose* est une modification linguistique consistant à ne pas prononcer un /h/ qui devrait être prononcé (ou une perte d'aspiration), par exemple en anglais. L'autre sens de ce mot est quand un

---

<sup>23</sup> Notre traduction

<sup>24</sup> <http://www.academie-francaise.fr/langue/questions.html#haricot>

<sup>25</sup> Wikipedia n'est pas la meilleure ressource en expertise linguistique mais représente tout de même une source d'information pour de nombreuses personnes et en ce sens représente une certaine norme et l'usage du peuple.

<sup>26</sup> [http://fr.wikipedia.org/wiki/Liaison\\_%28linguistique%29](http://fr.wikipedia.org/wiki/Liaison_%28linguistique%29)

locuteur fait en français une erreur avec un mot commençant par un h, que ce soit en réalisant l'élision ou la liaison avec un h aspiré, ou en traitant un mot avec h muet comme s'il avait un h aspiré<sup>27</sup>.

Spitzer (1941) explique l'origine du mot *pataquès* avec la fameuse histoire suivante :

Voici quelle en est, dit-on, l'origine : un plaisant était à côté de deux dames ; tout à coup il trouve sous sa main un éventail.

- Madame, dit-il à la première, cet éventail est-il à vous ?

- Il n'est point-z-à moi.

- Est-il à vous, Madame ? dit-il en le présentant à l'autre.

- Il n'est pas-t-à moi, Monsieur.

- Puisqu'il n'est point-z-à vous et qu'il n'est pas-t-à vous, ma fois, je ne sais pas-t-à qu'est-ce.

L'aventure fit du bruit et donna naissance à ce mot populaire, encore en usage aujourd'hui.

(Spitzer, 1941 : 253)

Spitzer (1941 : 253) précise que cette histoire est célèbre car citée par un grand nombre d'auteurs notamment dans les dictionnaires, manuels et ouvrages tels que Littré ou Vaugelas, qui relatent eux-mêmes l'histoire qu'on trouve dans *le Manuel des Amateurs de la langue* du grammairien Domergue (1805). Cette histoire est aussi connue car elle représente le goût des Français pour une langue bien parlée, mais aussi les critiques acerbes et les moqueries que provoquent les erreurs de liaisons.

Pour résumer, voici dans le tableau 1.22. les termes que la grammaire prescriptive attribue aux liaisons qui ne sont pas conformes aux règles du français normatif :

Terme	définition	Exemple
Liaison manquante	Non réalisation d'une liaison obligatoire	les / enfants
Pataquès	Insertion d'un s ou d'un t, Substitution d'un s pour un t	Ils devraient-z-être Pas-t-à moi, point-z-à moi
Cuir	Insertion d'un t	Il va-t-aller
Velours	Insertion d'un s	Quatre-z-yeux
Psilose	Liaison avec h aspiré Elision avec un h aspiré Liaison non faite avec h muet Non élision avec un h aspiré	Les z-héros L'héros Les/hôpitaux Le hôpital

Tableau 1. 22. Définitions des diverses erreurs de liaisons

<sup>27</sup> informations tirées du site : <http://monsu.desiderio.free.fr/curiosites/liaiserr.html>

Parmi ces erreurs de liaisons possibles, aucune étude ne mentionne des liaisons obligatoires (de type article + nom) manquantes dans la parole adulte francophone, ni même chez des enfants de plus de 6-7 ans, sauf dans des cas particuliers d'hésitations. Par exemple, dans leur étude du corpus PFC, Durand et Lyche (2008 : 53) n'ont répertorié que deux cas de liaisons obligatoires non réalisées avec le déterminant *les*, et les deux cas étaient avec des hésitations, lorsque les locuteurs changeaient d'avis et se reprenaient en changeant de structures (sans parler des cas qui empêchaient la liaison tels qu'avec des h aspirés).

Autrement dit, si erreurs il y a, celles-ci apparaissent dans des contextes de liaisons facultatives ou de liaisons interdites, c'est-à-dire dans des structures où les mots<sup>2</sup> sont rares ou très rares et lorsque la fréquence de cooccurrence du mot<sup>1</sup> et du mot<sup>2</sup> est faible, ou très faible. Il apparaît de tout ceci, que ce qui est considéré comme erreurs, c'est ce que les classes sociales qui ont le plus de pouvoir considèrent comme le moins acceptable, alors que celles que font les personnes plus éduquées, bien qu'aussi considérées comme fautives, sont plus acceptables. Autrement dit, la norme et le niveau « d'acceptabilité » sont dictés par le groupe de personnes ayant le plus de pouvoir : les personnes les plus éduquées.

#### **4.2. Cas particuliers: liaisons non-enchaînées et liaisons devant consonnes**

La règle générale concernant les liaisons est que si une consonne latente est prononcée, elle doit être enchaînée au mot<sup>2</sup> si celui-ci commence par une voyelle ou un h muet. Or, il existe des cas particuliers qui ne correspondent pas à cette règle générale. D'une part, on trouve quelques (rares) cas de liaisons avec consonnes. D'autre part, on trouve des cas (certes restreints) de francophones qui prononcent une consonne latente sans l'enchaîner, bien que cela aille à l'encontre du principe de syllabation ouverte.

On trouve d'abord les cas particuliers de *vingt* et *quand*. S'il est déjà particulier de prononcer une consonne latente avec une autre consonne, la prononciation de cette consonne latente est d'autant plus compliquée qu'elle dépend du contexte dans lequel *vingt* se trouve.

On prononce obligatoirement le *t* final de *vingt* devant ses dérivés (*vingt-deux*, *vingt-trois*, etc.) sauf si *vingt* est en fin de groupe rythmique (*j'en ai vingt*). Notons que, dans ce cas, le *t* est tout de même prononcé même en fin de groupe dans certains pays, notamment en Suisse. Par contre, on ne prononce pas le *t* s'il se trouve dans [quatre-vingt + consonne ou voyelle], comme par exemple *quatre-vingt-deux*, ou *quatre-vingt-huit*.

L'histoire de la prononciation de ce nombre est particulière (voir Malécot, 1954 : 290), et la prononciation des numéraux en général est problématique (liaison interdite avec *les un*, *les onze* ; le *f* se prononce /v/ dans *neuf heures* et *neuf ans*, et /f/ dans tous les autres cas, etc.). D'ailleurs, les nombres suscitent de nombreuses erreurs de fausses liaisons (*quatre-z-yeux*), tant chez les enfants que chez les adultes.

Il est aussi possible de prononcer le *t* de *quand*, même si celui-ci est suivi d'une consonne, comme le mentionnent Morin et Kaye (1982 : 300). Contrairement à *vingt*, ce n'est pas une obligation. Précisons que *quand* est souvent suivi de pronoms personnels (*il*, *ils*, *elle*, *elles*) et, dans ce cas, la liaison est considérée par certains (Léon, 1966 :129) comme obligatoire, pour d'autres (l'Académie Française) comme facultative (fréquente).

D'après Boula de Mareüil, Adda-Decker et Gendner (2003 :1330), la liaison est réalisée dans 94.6% des cas après *quand*, ce qui signifie qu'elle peut être considérée comme obligatoire ou tout au moins comme facultative très fréquente. Puisque la liaison est presque toujours produite, les locuteurs ont sans doute une représentation abstraite de *quand* comme ayant une consonne latente prononcée, voire comme étant une consonne fixe.

Comme l'observent Morin et Kaye (1982 : 300), on trouve des cas de liaisons avec *quand*, même si *quand* est suivi d'une consonne. Par contre, on ne trouve jamais le *d* de *quand* prononcé si *quand* se trouve à la fin d'une phrase (*je ne sais pas quand*). Puisque la consonne n'est jamais prononcée en fin de phrase mais seulement lorsque *quand* est suivi d'un autre mot, cela indique que dans le lexique des locuteurs, la consonne latente fait partie

de séquences dans lesquelles les mots<sup>2</sup> les plus fréquents commencent par une voyelle (*quand il, quand elle*). On peut supposer que les locuteurs anticipent la liaison avec la voyelle du mot<sup>2</sup> et généralisent cette liaison même aux cas où les mots<sup>2</sup> commencent par une consonne. La prononciation du d de *quand* pourrait alors représenter la fréquence d'occurrence dans laquelle *quand* apparaît avec des mots<sup>2</sup> commençant par une voyelle.

Il existe aussi les cas particuliers où certains francophones font des liaisons non-enchaînées pour des raisons stylistiques. Les liaisons non-enchaînées peuvent être fréquemment utilisées par l'élite française constituée de politiciens, journalistes, avocats et autres spécialistes de la parole (Encrevé, 1988), mais cette prononciation est considérée comme peu naturelle et pédante, par le reste de la population francophone. Durand et Lyche (2008 : 49) citent Passy (1892: 42), qui considère la « liaison non-enchaînée » comme ridicule ou risible, dans le passage suivant : « Et de fait, on cesse les liaisons dès qu'il y a arrêt. Rien de plus risible qu'une liaison faite mal a (sic) propos. *C'est une idée* prononcé (sɛt, ynide), fait croire qu'on a le hoquet. Un professeur prononçait des phrases comme *la première est excessivement facile* en s'interrompant après *est* ; (la pʁəmjeʁ ɛt ɛksesivmä fasil). La première fois que nous l'avons entendu, ça a été un éclat de rire général. »

Morin et Kaye (1982 : 299-300), en citant Ågren (1973 : 25), font état de cas de liaisons non-enchaînées devant consonnes, voyelles et coups de glottes. Notons que la plupart des liaisons non-enchaînées mentionnées par Morin et Kaye (1982) et Morin (2005 : 15) se trouvent avec la consonne /t/, c'est-à-dire dans les cas où, s'il y avait une voyelle après la consonne latente, ce serait une liaison facultative fréquente.

Morin (2005 : 15) remarque que les consonnes non-enchaînées devant consonne s'entendent fréquemment dans les médias et il en propose de nombreux exemples : « paraît [t] grave [...] est [t] pratiquement exclu [...] est [t] fort [...] C'était [t] juste avant [...] est [t] la conséquence [...] C'est [t] possible [...] sont [t] déterminés [...] sont [t] vraiment un

instrument [...] c'est [t] les [...] est [t] possible [...] sont [t] plus z-humiliés [...] Toute discrimination est [t] contre [...] Il faut [t] faciliter [...] Il était [t ?] hors [...] »

Notons que dans l'étude de Morin (2005 : 15), mis à part *paraît* et *faut*, toutes les autres liaisons non-enchaînées sont réalisées avec *être*. Or, la plupart du temps, *être* à la 3<sup>e</sup> personne du singulier est suivi de *un* ou de *une* (taux de cooccurrence de plus de 1000 pour les deux items), ce qui explique que ces séquences aient été enregistrées avec des liaisons dans ces contextes.

De plus, faire une liaison non-enchaînée permet à la fois de parler d'une façon prestigieuse (en faisant une liaison), mais aussi de parler de la façon particulière qu'utilisent les journalistes afin d'attirer l'attention des téléspectateurs, en utilisant l'accent d'insistance, placé sur la première syllabe.

Selon Gadet (1989), faire des liaisons non-enchaînées :

[...] permet en effet de combiner deux tendances partiellement contradictoires : effectuer la liaison (socialement valorisante), et faire porter l'accent d'insistance sur la première syllabe du mot :

(8) c'est « Impossible ([set èpösibl]

[...] Ce fait et sa fréquence en augmentation laissent entendre que l'enjeu social de la liaison l'emporte sur le respect du schéma canonique, puisqu'elle a pour effet d'augmenter le nombre de syllabes fermées. (Gadet, 1989 : 55)

D'ailleurs, il est prestigieux pour orateurs de faire des liaisons non-enchaînées (voir Encrevé, 1988). Ainsi, si des séquences se trouvant dans le lexique des locuteurs sont fréquentes, et que ces productions « hors normes », sont non seulement acceptées, mais mêmes prestigieuses, alors nous comprenons pourquoi il n'est pas surprenant de trouver ce genre de productions dans le contexte qui correspond à ce qui est attendu, dans un contexte de prestige ou le non-enchaînement des liaisons est prestigieux.

Toutefois les cas de liaisons non-enchaînées (ou les liaisons produites devant consonne) sont relativement rares en dehors des médias. Durand et Lyche (2008 : 53) ont observé que dans tout leur corpus de parole spontanée et de lecture de texte, seules trois liaisons ont été produites sans enchaînement avec *les*, et ceci était dû à des hésitations.

Le projet PFC fournit des résultats précis dans le pourcentage de liaisons enchaînées et non-enchaînées. Comme l'indiquent les tableaux 1.23. et 1.24. ci-dessous, en lecture, les liaisons (obligatoires et facultatives) sont réalisées et enchaînées dans 59.51% des cas, non réalisées dans 39.14% des cas, réalisées mais non-enchaînées dans 0.81% des cas. En parole spontanée, les liaisons (obligatoires et facultatives) sont réalisées et enchaînées dans 43.69% des cas, non réalisées dans 55.64% des cas, et réalisées mais non-enchaînées dans 0.25% des cas. Le taux de liaisons réalisées est plus élevé en lecture qu'en conversation.

lecture							
	LE	NR	NE	INC	EPE	HES	total
% liaisons	59.51	39.14	0.81	0.13	0.21	7.19	99.8

**Tableau 1. 23. % de réalisation de la liaison en lecture d'après les données du projet PFC**

LE : liaison enchaînée

NR : liaison non réalisée

NE : liaison non-enchaînée

INC : liaison incertaine (mauvaise qualité de l'enregistrement)

HES : hésitations (avec reprise ou non)

EPE : liaison épenthétique (fausse consonne de liaison)

conversation							
	LE	NR	NE	INC	EPE	HES	total
% liaisons	43.69	55.64	0.25	0.09	0.08	2.36	99.8

**Tableau 1. 24. % de réalisation de la liaison en conversation d'après les données du projet PFC**

Autrement dit, contrairement aux spécialistes de la parole qui utilisent cette technique pour des raisons stylistiques et expressives, les francophones ne font pas de liaisons non-enchaînées, sauf en cas d'hésitations qui sont souvent suivies de reprises. Quant aux liaisons « hors-norme », elles constituent une infime partie des productions. Soit elles sont réalisées selon la norme, soit elles ne sont pas réalisées du tout dans la majorité des cas.

Mallet (2008 :178-183) rapporte les résultats observés d'après le même corpus PFC en ce qui concerne les liaisons non-enchaînées, qu'elle relève au nombre de 130. Elle en vient à la même conclusion que Durand et Lyche (2008), c'est-à-dire qu'il s'agit d'erreurs de performance et non de véritables cas de liaisons non-enchaînées au sens d'Encrevé :

[...] ce qui se révèle particulièrement caractéristique est la présence de fortes hésitations. Ces hésitations sont attestées par les nombreuses reprises opérées par les locuteurs [...] Les liaisons non-enchaînées sont souvent marquées par des interjections (euh) et il y a d'autres signes d'hésitations, notamment les reprises. [...] D'après nos estimations, 20 à 30 % des liaisons identifiées au départ comme « non-enchaînées » dans les conversations libres et guidées ne constituent pas de « vrais » cas de non-enchaînement au sens d'Encrevé. (Mallet, 2008 : 178-183)

Mallet (2008 : 178) dénombre 130 liaisons non-enchaînées et explique (2008 : 183) que parmi les 48 liaisons épenthétiques il y avait en fait : « 28 erreurs de codage, 6 liaisons avec consonne différente, 14 épenthèses au sens de PFC (fausse liaison), dont 13 en lecture (qui mène-t-au village). » Autrement dit, les liaisons non-enchaînées sont dans ce corpus des erreurs de parcours (ou de codage) car elles ne sont ni systématiques, ni séparées de l'hésitation, ni une véritable volonté de non-enchaîner les consonnes de liaison, contrairement aux politiciens étudiés par Encrevé (1988), qui choisissent de ne pas enchaîner les liaisons pour des raisons stylistiques.

Durand et Lyche (2008 : 49) signalent en se référant aux travaux de Laks (1983) que soit les jeunes des banlieues réalisent les liaisons facultatives enchaînées, soit ils ne les réalisent pas du tout. Par contre, ils ne les réalisent jamais non-enchaînées. On peut donc en conclure que la liaison non-enchaînée, qu'elle soit devant une voyelle, devant une consonne, avec un coup de glotte ou non, n'est utilisée que par une très faible partie de la population francophone à des fins stylistiques dans le contexte très particulier des médias. Sinon, il s'agit d'hésitations suivies le plus souvent de reprises et non d'un véritable non-enchaînement.

Nous reparlerons des liaisons non-enchaînées lorsque nous analyserons les productions de francophones et d'anglophones dans notre étude.

### **4.3. Les erreurs de liaisons : explications**

Les erreurs nous informent sur le traitement des informations concernant la liaison. Elles nous informent sur ce qui est « naturel » pour les francophones, même si cela va à l'encontre de la norme. Si la contrainte anti-hiatus (ou anti attaque vide) peut expliquer certaines erreurs que font les enfants et certains adultes en rajoutant des consonnes épenthétiques (*quatre z-yeux*), de nombreux phénomènes soit n'ont pas pour origine cette contrainte, soit vont même à l'encontre de cette contrainte.



Les enfants apprennent les mots avec leurs liaisons, comme si la consonne de liaison faisait partie du mot<sup>2</sup>, comme si ces consonnes étaient des proto-déterminants (Wauquier-Gravelines et Braud, 2005), ce qui n'est pas sans rappeler Gougenheim (1938) qui avait déjà émis l'hypothèse que la consonne de liaison en [z] pouvait être une sorte de préfixe [z] du pluriel des noms et adjectifs.

Morin (2005) relate les explications d'un francophone de Louisiane, scolarisé en anglais et ne sachant ni lire ni écrire le français. Les exemples donnés dans cet extrait sont tout à fait similaires aux productions d'enfants d'âge pré-scolaire et sont représentatifs du processus d'acquisition :

— Comment vous dites cold ?

— [lœfrɛ]. C'est-à-dire, c'est selon l'histoire [lœ distwar] /c'-à-d. le contexte/. Tu vois, pour un n-exemple, t'as larbre, narbre, arbre ou zarbre : un narbre. Tu vois, t'as /hésite et se reprend/ des fois t'uses le mot larbre, narbre, arbre ou zarbre. Zarbre veut dire « plus [plys] qu'un ». En anglais, t'uses un mot. Ça me gêne pas si y en a un ou i n'n a dix, c'est toujours le même mot. Et en français, t'as quatre mots pour un narbre ou un n... /hésite/ un /se reprend/ larbre. C'est selon, comment t'appelles /hésite et se reprend/ dépeins ton discours, là tu uses le mot zarbre, parce que tu peux dire quand quelqu'un n'est pas habitué de parler en français un tas, il va dire Regarde le gros narbre. Non, non. Tu dis : Regarde le gro arbre [gro a:b], comprends ? (Morin, 2005 : 14)

Si un francophone se trouve dans un milieu où le français est une langue minoritaire, comme c'est le cas des francophones de Louisiane, où l'input a été restreint pendant la période de son apprentissage, et où, de surcroît, le français connaît l'interférence de la langue environnante (l'anglais), la langue sera différente (et la production de liaison aussi) de celle des francophones majoritaires.

Si, en plus de tout cela, le francophone n'a jamais vu les séquences écrites de ces liaisons, ses connaissances du fonctionnement de la liaison sont, d'une certaine façon, comparables à celles de jeunes francophones de moins de 6 ans qui ne savent pas encore lire et écrire. Le locuteur fera alors uniquement les liaisons qui sont les plus ancrées dans son lexique, d'après la productivité et la fréquence des structures et des items. Il fera donc les liaisons les plus fréquentes souvent entendues pour être mémorisées, fera peu de liaisons moins fréquentes (liaisons facultatives), ou les fera avec des erreurs.

Les productions du francophone adulte de Louisiane mentionné dans Morin (2005 : 14) sont similaires aux productions d'enfants relevées par exemple dans Dugua (2005 et 2006) où elle aborde notamment les erreurs de remplacement de type *le navion* ou *un zavion*. L'expérience langagière du locuteur de Louisiane a sans doute été moins riche que celle d'un jeune locuteur de langue majoritaire car il aura entendu moins de séquences conformes aux règles de la langue normative et une plus faible proportion de liaisons facultatives que le jeune locuteur francophone. Sa perception et ses productions de liaisons sont directement liées à son input langagier et à ses conditions d'apprentissage.

Mis à part certains professeurs de FLE/FL2, les francophones ne connaissent généralement pas les règles théoriques concernant les liaisons (ce qui a été d'ailleurs été confirmé dans les réponses à notre questionnaire). Les locuteurs natifs connaissent les règles phonétiques par intuition, en ayant entendu des structures fréquentes qu'ils ont mémorisées.

Toutefois, s'ils ne peuvent pas se baser sur des règles théoriques, et que leur intuition dépend de l'input de leur entourage, leurs connaissances sont forcément limitées par leur expérience langagière. Des lacunes dans l'apprentissage et un input limité dans certains domaines (les liaisons facultatives plus rares) vont inévitablement mener les locuteurs à commettre des erreurs.

Si les règles phonétiques sont enseignées, c'est généralement à l'école maternelle ou à l'école primaire, en même temps que les francophones apprennent à lire et à écrire. C'est ce que les enseignants nomment « conscience phonologique ». Les jeunes apprenants dans des écoles francophones, notamment en cours préparatoire en France (ce qui correspond à la 1<sup>ère</sup> année au Canada) de l'étude de Sprenger-Charolles, Siegel et Bonnet (1998) se basent, au début de leur apprentissage de la lecture, sur leurs connaissances phonologiques pour apprendre à lire et à écrire en français. Ils voient un mot écrit et essaient de trouver à quel mot entendu le mot écrit correspond. Ils lisent à haute voix un mot, jusqu'à ce qu'ils trouvent un

mot qui, lu à haute voix, correspond à un mot existant à l'oral dans leur lexique.

Puis, au fur et à mesure qu'ils rencontrent des mots en lecture, ils les reconnaissent. Ils ont alors de moins en moins besoin de se référer à l'oral, et se basent davantage sur leurs connaissances de l'écrit en faisant des analogies avec d'autres mots écrits. Ils continuent néanmoins à utiliser « la médiation phonologique », mais moins souvent qu'en début d'apprentissage.

Les connaissances phonétiques intuitives sont indispensables pour toutes les activités orales, que ce soit la parole spontanée ou la lecture. En outre, la fréquence des mots écrits joue un rôle dans la reconnaissance des mots et de leur lecture (Ellis 2002 :150-152). Cela signifie donc que, si les enfants n'ont pas entendu les mots dans leur entourage et ne les ont pas assez rencontrés à l'écrit, il est fort probable qu'ils fassent des erreurs de lecture.

De plus, les règles sont souvent complexes. Il suffit de prendre comme exemple, les treize pages d'explications et de listes relatives aux liaisons interdites et aux h aspirés de Fouché (1959: 252-265). En outre, les règles ne sont pas « absolues » et les classifications (liaisons obligatoires vs liaisons facultatives) varient selon les auteurs. Enfin, il existe un flottement pour certaines règles et pour certains mots commençant par un h ou une semi-consonne (et d'autant plus pour des mots commençant par un h et une semi-consonne, comme c'est le cas pour le mot *hiatus*).

Par exemple Gaillard et Collignon (2005 : 75) expliquent qu'il y a un flottement dans la règle des h dans les noms propres : « De Henri IV » ou « d'Henri IV », « d'Hugo » ou bien « de Hugo » ?... Hélas, ou tant mieux, c'est *ad libitum* : les deux traitements sont usités et donc avalisés. Mêmes tolérances pour des prénoms comme Henriette, Hugues, Hubert... Mais personne, en France, n'ira dire ou écrire : « La fille d'Henriette s'installe à L'Havre », au lieu de « au [à Le] Havre ».

En fait, on voit ici qu'au départ il y avait une règle (de ne pas faire l'élision ou la

liaison), mais que lorsque les mots sont devenus courants (*Henri, Henriette, Hugo*), la fréquence a permis au locuteur de mémoriser ces mots et séquences, en changeant le traitement.

Précisons également que dans un même contexte syntaxique les règles ne sont pas toujours les mêmes, car elles dépendent aussi d'autres facteurs, tels que des facteurs phonétiques. Par exemple, la liaison dans le contexte [adj. + nom], dépend de plusieurs facteurs autres que syntaxiques ou phonétiques.

Selon Morin et Kaye (1982 : 301 et 317), toutes les consonnes de liaisons n'ont pas le même statut. Tous les cas liants avant une pause (d'hésitation ou d'intonation) ou avant un coup de glotte excluent la liaison avec *n* et la liaison avec *z* après des déterminants et des noms pluriels. Selon eux, le fait qu'il puisse y avoir une liaison après une pause indique que cela n'est pas nécessairement une marque de cohésion phonétique forte entre deux mots lexicaux. Contrairement aux idées de Selkirk (1974), la liaison est possible entre une tête lexicale et son complément, qu'il soit étroitement lié au verbe ou non (*ils arrivent\_en courant*)

De même, on voit l'influence du facteur phonétique, dans les catégories syntaxiques, puisque la liaison qui était avant possible avec les consonnes /t/ (*un soldat t-anglais*), /p/ (*du drap p-anglais*) et /k / (*un sang k-impur*), n'a jamais été possible avec /n/ (*\*un lapin n-anglais, \*un garçon n-intelligent*), pourtant dans le même contexte syntaxique [nom + adj. sing.] (Morin et Kaye, 1982 : 317)

Enfin, les liaisons avec /n/ sont complexes, car il peut y avoir trois cas de liaisons dans le contexte [adj. sing. + nom] : liaison avec voyelle nasale, liaison avec voyelle dénasalisée et pas de liaisons, comme le remarque Sampson (2001). Autrement dit, pour maîtriser les liaisons, il faut connaître (ne serait-ce que de façon intuitive) au minimum les règles syntaxiques et phonétiques.

Il est quasi impossible pour des francophones de se baser sur des règles théoriques, car

celles-ci sont nombreuses, complexes et peu ou pas étudiées. Ils ont donc des connaissances intuitives, mais celles-ci dépendent principalement de leur milieu linguistique, et donc de leur statut socio-économique. Ils prononcent les séquences, telles qu'ils les ont entendues dans leur entourage, d'après ce qui leur semble juste. C'est pourquoi leur production de liaisons facultatives plus rares dépend en grande partie de facteurs sociolinguistiques.

S'ils viennent de milieu défavorisé, ils n'auront pas pu entendre ces liaisons assez souvent dans leur entourage pour les maîtriser (Chevrot, Dugua et Fayol, 2005 ; Pierrehumbert, 2002 ; Chevrot, Chabanal et Dugua, 2007, par exemple). S'ils n'ont pas bénéficié d'une exposition à un nombre suffisamment élevé d'occurrences pour apprendre des séquences de façon intuitive et n'ont pas appris les règles de façon formelle, les séquences moins fréquentes telles que les liaisons facultatives et les liaisons interdites vont créer des hésitations et des erreurs, et cela est d'autant plus le cas dans des situations formelles où l'on peut faire de l'hypercorrection.

Les liaisons sont problématiques car si un locuteur natif n'en fait pas assez, comme le disent Green et Hintze (1990 : 83) : « [...] in a highly literate culture, fear of being thought “uneducated” is a powerful deterrent ». S'il fait des erreurs, cela le rend ridicule. Et s'il fait plus de liaisons que nécessaire, il paraîtra snob ou pédant. Dès qu'un locuteur francophone se détache de la norme, il est jugé négativement, comme le remarquent Peyrollaz et de Tovar (1954 : 177) : « [...] les fautes de liaisons sont particulièrement choquantes en français, où elles dénotent une prononciation provinciale ou étrangère, voire pédante et prétentieuse.[...] »

Les francophones savent, par intuition, que produire plus de liaisons paraît plus « éduqué ». Par contre, puisqu'ils ne connaissent pas explicitement les règles de liaison, surtout pour les liaisons facultatives et les liaisons interdites, ils vont parfois commettre des erreurs, ce qui, comme le pense Spitzer (1941 : 254) trahit leurs origines sociales : « [...] les enrichis sans éducation se trahissent par leurs pataqui-pataquiès ».

Les fausses liaisons donnent une mauvaise impression, car si les locuteurs n'essayaient pas de paraître « éduqués », ils ne feraient pas de liaisons facultatives (incertaines) du tout, et ne feraient donc pas ces erreurs, comme c'est le cas par exemple avec les personnes peu éduquées mentionnées dans Morin (2005 :12), qui relate l'étude de Durand (1936 : 241) : « Lorsqu'ils sont forcés de produire des constructions délicates (à la lecture par exemple), la stratégie la plus fréquente consiste à ne pas faire la liaison du tout. [...] ».

Dans des situations d'insécurité linguistique et d'hypercorrection, la peur de paraître ignare pousse les locuteurs doutant de leurs connaissances linguistiques à prendre des risques et à produire des liaisons dans des séquences qu'ils n'ont pas assez entendues pour les maîtriser. Ils réalisent alors une liaison, par analogie, d'après les séquences les plus proches qui leurs viennent à l'esprit.

Les erreurs de langue sont souvent condamnées et relèguent le locuteur au rang de personne inculte. D'où les nombreux organismes ou gardiens de la langue qui veillent au maintien du bon usage et qui mettent en garde les locuteurs contre les effets négatifs des erreurs de prononciation, comme par exemple dans l'extrait suivant de Wikipedia :

Certains termes français à h aspiré sont connus, par leur fréquence et leur caractère commun, comme haricot ou hérisson. Ne pas prononcer la disjonction avec ces mots est considéré comme particulièrement populaire. D'autres, cependant, sont régulièrement prononcés sans leur h car c'est la majorité des locuteurs qui pense les prononcer correctement : handicaper et tous ses dérivés débute par un h aspiré. Il est cependant rare de l'entendre prononcer, d'autant plus à la radio ou la télévision, source linguistique d'une grande partie de la population. Enfin, par hypercorrection, on peut ajouter un h là où il n'y en a pas : beaucoup prononcent un hiatus avec un tel h, et écrivent le hiatus là où l'hiatus est attendu. De même, certains hellénistes disent et écrivent le hoplite au lieu de l'hoplite, par contamination avec la connaissance du grec ancien<sup>28</sup>.

Les francophones acceptent d'autant moins les erreurs de prononciation (ici les liaisons) et les trouvent d'autant plus impardonnables qu'elles sont présentes dans les médias, comme en témoignent les nombreuses plaintes des lecteurs et téléspectateurs notamment adressées au CSA (Conseil Supérieur de l'Audiovisuel) :

---

<sup>28</sup> <http://fr.wikipedia.org/wiki/Psilose>

Les liaisons font partie des pièges les plus courants de la langue parlée. Elles suscitent de nombreuses lettres de téléspectateurs et d'auditeurs courroucés qui demandent au CSA de « faire en sorte qu'enfin on cesse de prendre tous les Français pour des analphabètes, quand on a le privilège de s'adresser à plusieurs millions de personnes.<sup>29</sup>

Ceci est peut-être le cas, parce que traditionnellement les personnes travaillant dans les médias sont les spécialistes de la parole, et celles qui utilisent le plus les liaisons non-enchaînées, signe de prestige et d'éducation. Les auditeurs s'attendent donc à une langue de qualité « supérieure ».

Toutefois, si les liaisons sont déjà un problème, même pour les spécialistes de la parole, elles le sont d'autant plus pour les non-spécialistes qui sont tout de même jugés lorsqu'ils se trouvent en situation orale. Ces erreurs de liaisons représentent un problème si important pour les Français qu'Esposito (1994) en a fait un livre de 93 pages utilisé pour la préparation de concours.

En outre, certains Français se plaisent à souligner et critiquer les erreurs linguistiques des autres, d'où de nombreux sites web qui en font la liste<sup>30</sup>. Peut-être que se moquer des erreurs des autres est une façon de revendiquer leur statut socioéconomique puisque seules les personnes éduquées connaissent et surtout maîtrisent les difficiles règles du français, et encore bien moins peuvent affirmer sans aucun doute qu'il s'agit d'erreurs. Peut-être que les structures fautives vont à l'encontre de la grammaire des locuteurs, de leurs connaissances et intuitions, et en ce sens, cela les dérange.

Sur le site mentionné en pied de page, où les internautes envoient des textes et vidéos notant les erreurs des auteurs, on trouve par exemple :

"dans les 20 **z**-arrondissements de Paris" [...] "avec plus de 100 **z**-idées gourmandes... le conseil de 100 **z**-idées au jardin" [...] "les 80 **t**-autres pourcents" [...] "il est donc **t**-inutile de préciser" [...] "main qui, selon **t**-eux" [...] "je vis **n**-à Neuilly" [...].

---

<sup>29</sup> [http://www.csa.fr/actualite/dossiers/dossiers\\_detail.php?id=129511](http://www.csa.fr/actualite/dossiers/dossiers_detail.php?id=129511)

<sup>30</sup> <http://www.retourneaucm1.com/index.php?2008/01/13/212-une-montagne-de-cuir-et-de-velours>

Notons d'ailleurs que la plupart des erreurs surviennent avec les numéraux. Paradoxalement, alors que certaines sont sévèrement condamnées, d'autres s'entendent régulièrement dans la conversation de tous les jours, et sont plus ou moins acceptées à l'oral (*quatre-z-yeux, donne-moi-z-en*), tout comme on en trouve dans les chansons populaires, notamment dans la chanson enfantine (du XVIII<sup>e</sup> siècle) *Malbrouk s'en va-t-en guerre* :

Malbrouk s'en va-**t**-en guerre  
Mironton tonton mirontaine  
Malbrouk s'en va-**t**-en guerre  
Ne sais quand reviendra (bis)  
Il reviendra-**z**-à Pâques  
Mironton tonton mirontaine [...]  
J' l'ai vu porté en terre/  
Par quatre-**z**-officiers

On retrouve d'ailleurs le même phénomène dans la chanson enfantine *A la claire fontaine* (qui remonterait au XII<sup>e</sup> siècle) : « Tu as le cœur à rire, moi je l'ai-**z**-à pleurer ».

Pourtant, comme le constate Lauret (2007 : 17), dans l'ensemble les francophones sont peu indulgents envers les erreurs langagières commises par des étrangers et c'est particulièrement le cas pour la prononciation : « [...] Les Français et les Japonais sont décrits comme particulièrement sensibles et exigeants face aux étrangers qui parlent leur langue [...] ».

Les francophones jugent ceux qui parlent le français (les francophones et les non-francophones) d'après des règles normatives de prononciation. Les accents sortant de la norme, tels que les accents régionaux, tout comme les erreurs de langue en général, sont souvent perçus comme étant révélateurs du milieu socioéconomique du locuteur, et les accents étrangers peuvent produire des impressions négatives qui s'étendent au delà de la perception phonétique.

Dans l'étude de Tardif et d'Anglejan (1981 : 210), des sujets ont évalué selon plusieurs critères la qualité de la langue de deux locuteurs lisant des phrases. Le but de l'étude était de tester l'impact de la prononciation d'un locuteur sur la perception globale de son



français. La moitié des juges a évalué des phrases lues avec un accent anglophone très marqué (voix n° 1), l'autre moitié a évalué les mêmes phrases lues avec un accent anglophone peu marqué (voix n° 2) : « [...] de façon globale, les juges natifs évaluent les phrases lues par la voix N° 1 comme étant moins compréhensibles [...], moins exactes grammaticalement [...], moins admissibles [...], et plus irritantes [...] que les mêmes phrases lues par la voix N° 2 ».

La prononciation d'un locuteur influence le jugement des interlocuteurs francophones non seulement quant aux compétences phonétiques de celui-ci, mais aussi quant aux autres aspects linguistiques tels que la grammaire et le vocabulaire. Alors que deux sujets lisaient exactement les mêmes phrases avec un accent plus marqué que l'autre, la perception de l'accent influence la perception de la maîtrise de la langue en général, au-delà de la prononciation. Pour des francophones, les erreurs de prononciation sont perçues comme étant du domaine de la langue en général et de la grammaire.

On retrouve un phénomène similaire dans le sketch des Nuls, *Les liaisons dangereuses*<sup>31</sup>. Le passage contient quinze fautes de liaisons (liaisons interdites réalisées et surtout cuirs, velours et pataquès) et six de grammaire. Le commentateur de la soi-disant bande annonce du film conclut en disant « un film époustouflant avec que des fautes d'accord ».

Pour des francophones qui ne connaissent pas les règles, les fautes de liaison sont perçues comme des fautes de grammaire, et plus précisément des fautes d'accord. Il est vrai qu'il y a eu six fautes de grammaire dans ce sketch, mais, parmi toutes ces fautes, il n'y en a véritablement eu que deux d'accord sur un total de 21 fautes, dont 15 de liaisons. Les acteurs veulent paraître éduqués et pressentent que les liaisons ajoutent de la distinction, mais ils n'ont pas toujours les connaissances linguistiques nécessaires pour les faire correctement.

---

<sup>31</sup> [http://www.dailymotion.com/video/xtplh\\_les-liaisons-dangereuses\\_fun](http://www.dailymotion.com/video/xtplh_les-liaisons-dangereuses_fun)

Dans les deux cas, dans l'étude de Tardif et d'Anglejan (1981) et dans le sketch des Nuls, les sujets ont du mal à séparer des erreurs de prononciation d'autres composantes linguistiques, notamment syntaxiques. Les interlocuteurs remarquent que la production est différente de la norme attendue, mais ils ne peuvent pas identifier avec précision les erreurs et séparer les composantes linguistiques.

Une approche selon laquelle le lexique est au centre de toute activité linguistique explique en partie ce comportement. D'après le MBU, les jeunes locuteurs apprennent des mots ou des séquences de mots et c'est en les ayant mémorisés et en ayant constitué une banque de données permettant des généralisations abstraites qu'ils peuvent maîtriser des structures et des items et en déterminer leur validité et leur grammaticalité.

Selon le MBU, et en particulier Ellis (2002) et Tomasello (2000a et 2000b), la langue (son acquisition et sa production) est basée sur des mots, tout est lié, notamment la morpho-syntaxe et la phonétique. Boula de Mareuil, Adda-Decker et Gendner (2003 : 1332) concluent aussi à la fin de leur article en faisant référence à Fougeron, Goldman et Frauenfelder (2001) « This confirms that liaison is strongly linked to both syntactic and lexical levels. »

C'est la raison pour laquelle les locuteurs ont des difficultés à séparer la langue en sous-catégories, car lexique, phonétique et morphosyntaxe sont intimement liés. Les locuteurs ont une intuition de ce qui est grammatical et de ce qui ne l'est pas, sans vraiment savoir pourquoi une phrase est correcte ou non. Leurs intuitions de natifs se basent sur une mémoire de ce qui a suffisamment été entendu pour faire partie des structures abstraites et pour déterminer la précision d'une phrase et sa grammaticalité.

Toutefois, si de nombreuses composantes linguistiques demeurent entièrement intuitives car non étudiées formellement (la phonétique), les locuteurs connaissent mieux une composante que toutes les autres : la morphosyntaxe. En raison de leur formation scolaire, la

grammaire et l'orthographe sont probablement ce que les francophones connaissent le mieux (comparé à la phonétique). Laks (2005b) rappelle l'importance de l'apprentissage de l'écrit :

L'écrit se généralise [...] et impose sa légitimité jusqu'à contraindre les usages oraux. [...] la scolarité obligatoire couvre au minimum dix à douze des premières années de la vie et [...] l'apprentissage explicite de l'orthographe y est central. Quotidiennement exercé et renforcé, l'enseignement de la grammaire scolaire, avec ses manipulations de listes de mots, de règles et d'exceptions et ses raisonnements paradigmatiques exemplaires, est au cœur du dispositif pédagogique.  
(Laks, 2005b : 114)

L'oral, et notamment les liaisons, a un rapport direct avec l'orthographe des mots et les règles de grammaire correspondantes. Par exemple, connaître les mots qui ont un h aspiré peut être considéré comme faisant partie des règles grammaticales, car cela déterminera notamment les choix d'article (*le* vs *l'*) et d'adjectif (*ce* vs *cet*, *nouveau* vs *nouvel*, etc.) et ce type de règles se trouve dans les manuels de grammaire et d'orthographe, comme dans le livre de Gaillard et Colignon (2005 : 75), *Toute l'orthographe*.

De plus, savoir s'il y a un *s* ou non à certains mots tels que les numéraux, demande non seulement de connaître leurs règles de grammaire d'accord (savoir si *cent*, *vingt*, *mille* prennent un *s* ou non), mais aussi d'être capable d'appliquer ces règles en quelques millisecondes. De simples connaissances théoriques ne suffiraient pas à produire rapidement des séquences bien formées. Autrement dit, il faut connaître intuitivement les règles de grammaire et être capable de les utiliser de façon automatique, sans réfléchir.

Par exemple, pour ne pas faire d'erreurs avec les numéraux, il faut parfaitement connaître les règles d'accord et avoir entendu et produit ces séquences de nombreuses fois. Vu le temps limité de réflexion à l'oral, le locuteur utilisera la séquence la plus souvent entendue et pratiquée. Or, la plupart du temps, les expressions du pluriel sont *les*, *des* ou *deux*, ou à la limite *trois*. Au-delà de *quatre*, les cas d'utilisation des chiffres sont beaucoup plus rares.

Desrochers (1994) a étudié les erreurs de liaison pour déterminer leurs contextes et leurs origines et il en a déduit qu'il y a plusieurs causes selon le type d'erreurs : lexicalisation, morphologisation et préfixation et parfois même plusieurs phénomènes à la fois.

Contrairement à Desrocher qui considère que les erreurs ont des origines différentes, nous soutenons qu'elles ont toutes la même origine : la fréquence de séquences et de structures.

En utilisant la banque de données de Lexique<sup>32</sup> ( Base, Film2-Tout), si l'on considère le taux de cooccurrence de quelques expressions au pluriel, on obtient les taux de cooccurrence suivants :

"des amis" = + de 1000 ; "les amis" = 999 ; "deux amis" = 109 ; "trois amis" = 11 ;  
"quatre amis" = 2 ; "cinq amis" = 2 ; "huit amis" = 0 ; "vingt amis" = 0.  
"des autres" = + de 1000 ; "les autres" = + de 1000 ; "deux autres" = 709 ; "trois autres" = 203 ;  
"quatre autres" = 60 ; "cinq autres" = 35 ; cent autres" = 10 ; "mille autres" = 15 ;

Nous voyons donc que les expressions du pluriel sont en grande majorité utilisées avec des mots permettant (ou imposant) une liaison avec /z/ (*des, les, deux, trois, six, dix*) et sont de plus en plus rares au fur et à mesure que les chiffres augmentent (*quatre, cinq, huit, vingt, cent, mille*). Les locuteurs entendent donc beaucoup plus de séquences avec /z/ que sans /z/ et de séquences contenant des chiffres inférieurs à 3 plutôt que supérieurs à 3. Cela explique des erreurs du type : « *quatre-z-enfants* » (Morin, 2005 :11).

Il en est de même avec certains préfixes. Les préfixes sont invariables et ne prennent pas le s du pluriel. C'est le cas par exemple de *des symptômes gastro-z-intestinaux* (Morin et Al, 1990 : 524). Le problème est que toutes les séquences impliquant le mot *intestinaux* sont rares :

symptômes intestinaux : fréquence de cooccurrence = 0;  
problèmes intestinaux : fréquence de cooccurrence = 4;  
gastro-intestinaux = 4

Il y a ici un conflit qui entraîne des erreurs. D'une part, la plupart des francophones savent que *gastro* ne prend pas de *s*, tout comme ils savent qu'il n'y a pas de *s* à *quatre* de *quatre-z-yeux*. D'autre part, les francophones savent que la plupart des séquences au pluriel prennent le /z/ de liaison. Ici, un taux de cooccurrence très bas pour toute les séquences possibles avec *intestinaux* renforce ce conflit. Le locuteur utilise la séquence la plus

---

<sup>32</sup> <http://www.lexique.org/moteur/RechCoOc.php>

productive et surgénéralise la séquence la plus fréquente, qui est le /z/ du pluriel. Ainsi, la productivité de structures et la mémoire de séquences entendues (avec /z/) semblent être plus puissantes que la connaissance de règles grammaticales et/ou phonétiques, ou l'influence de l'écrit.

Les compétences langagières (compréhension et production) sont basées sur une analyse statistique qui détermine les probabilités de production de structures et séquences les plus fréquentes (voir par exemple Saffran et al, 1996 ; Juliano et Tanenhaus, 1993). D'autre part, les éléments sont codifiés dans un système hiérarchique. Les éléments les plus entendus se trouvent en haut de la hiérarchie et sont le plus facilement accessibles d'après le « Modèle de la Concurrence » (Mac Whinney, 1978, 1982, 1987).

Les erreurs à distance sont représentatives du processus de liaison dans le cadre du MBU. Les locuteurs ont accès dans leur lexique à une banque de données, dont les séquences sont traitées selon plusieurs paramètres dont la fréquence. Les constructions les plus fréquentes et les plus productives se trouvent en haut de la hiérarchie. À la lumière du MBU, nous pouvons comprendre les causes des erreurs citées dans Morin et Kaye (1982).

- (78) (a) Il devra-t-y avoir [...]
  - (b) les résultats qu'on aurait pu-t-en espérer [...]
  - (c) J'ai beaucoup-t-aimé [...]
  - (d) ça doit bien -t-être cuit [...]
  - (e) ils ne veulent pas-t-aller à l'école [...]
  - (79) (a) C'est l'image qui m'est venue-t-à l'esprit. [...]
  - (b) Ils chantent tous-t-en cœur [...]
  - (c) C'est pas-t-à moi [...]
  - (d) Je ne sais pas-t-à qui c'est [...]
  - (e) Donne-moi-t-un conseil [...]
  - (f) Laisse-moi-t-ici [...]
- (Morin et Kaye, 1982 : 323-324)

Nous avons vu en expliquant le MBU et la productivité des structures que, d'après le Modèle de la Concurrence développé par MacWhinney (1978, 1982, 1987) et Bates et MacWhinney (1979, 1982), les structures et les séquences les plus fréquentes sont accessibles en priorité, tandis que les séquences et éléments moins fréquents se trouvent en bas de la

hiérarchie. Si l'on considère les exemples de Morin et Kaye (1982 : 323-324), on peut imaginer (dans la colonne de droite) les structures qui se trouvent en haut de la hiérarchie (tableau 1.25.) et ainsi mieux comprendre les causes de ces erreurs.

Erreurs mentionnées dans Morin et Kaye (1982 : 323-324)	Structure en haut de la hiérarchie
(a) Il devra-t-y avoir [...]	il doit / devait / devrait y avoir
(b) les résultats qu'on aurait pu-t-en espérer [...]	on peut / pouvait/ pourrait espérer
(c) J'ai beaucoup-t-aimé [...]	J'ai vraiment /surtout, ils sont / aimé
(d) Ça doit bien -t-être cuit [...]	ça doit / devait / devrait être cuit
(e) ils ne veulent pas-t-aller à l'école [...]	ils veulent aller à l'école
(a) C'est l'image qui m'est venue-t-à l'esprit. [...]	C'est l'image qui vient à l'esprit
(b) Ils chantent tous-t-en cœur [...]	Ils chantent en cœur
(c) C'est pas-t-à moi [...]	C'est à moi
(d) Je ne sais pas-t-à qui c'est [...]	C'est à qui ?
(e) Donne-moi-t-un conseil [...]	Ils donnent un conseil
(f) Laisse-moi-t-ici [...]	Laissent ici

**Tableau 1. 25. Notre interprétation des séquences se trouvant en haut « de la hiérarchie » expliquant les erreurs d'après les exemples de Morin et Kaye (1982 : 323-324)**

Nous voyons que les structures de la colonne de droite du tableau 1.26. sont plus fréquentes que celles de la colonne de gauche. Les locuteurs ont dans leur lexique les séquences les plus fréquentes (de droite), et s'ils changent légèrement ces structures en y ajoutant un adverbe, en utilisant l'inversion ou en changeant de personne, alors les séquences restent « intactes », telles qu'elles sont dans leur lexique produisant alors une inadéquation avec les éléments moins fréquents qui sont alors des « intrus ».

Il en est de même avec les exemples de Gadet (1989) (tableau1.26.). Les séquences beaucoup plus fréquentes (dans la colonne de droite) dans le lexique des locuteurs sont accessibles avant les séquences moins fréquentes (colonne de gauche) créant ainsi des erreurs.

Erreurs mentionnées dans Gadet (1989 : 56 )	Structures correspondantes en haut de la hiérarchie
(13) moi z-aussi [...]	vous aussi
(14) les auteurs et les non z-auteurs [...]	les auteurs
(15) les cinq z-années à venir [...]	les années / les dix années
(19) douze z-invités [...]	des invités / les invités / deux invités / trois invités
(20) si vous laissez r-un message [...]	(veuillez) laisser un message

**Tableau 1. 26. Notre interprétation des séquences se trouvant en haut « de la hiérarchie » expliquant les erreurs d'après les exemples de Gadet (1989 : 56 )**

On voit dans la colonne de droite ce qui est en haut de la hiérarchie du locuteur, ce qui a causé les erreurs de la colonne de gauche, ce qui reflète le lexique du locuteur, l'input reçu et la fréquence de certaines structures par rapport à d'autres.

Toutes ces erreurs ont la même origine : les représentations abstraites du locuteur sont basées sur la saillance, la récence et surtout la fréquence de certaines structures qui l'emportent sur des séquences ou des structures moins fréquentes. Le MBU propose une approche qui unifie les différents types d'erreurs des enfants et des adultes et explique de nombreux phénomènes. La position, l'ordre et la prononciation de ces séquences dépendent de la fréquence de la structure, du mot1, du mot2, de la fréquence de cooccurrence du mot1 et du mot2 et de la fréquence du phonème de liaison et d'un système basé sur une approche lexicale probabiliste qui participe à la production de liaisons.

## **Discussion**

Nous avons vu dans ce chapitre plusieurs phénomènes qui se regroupent et ont la même origine. La production de liaisons facultatives des adultes dépend de leur éducation, de leur milieu social et des structures et séquences dans lesquelles se trouvent les liaisons. Les structures et les items plus rares sont traités comme les structures et les items plus fréquents, par effet d'analogie et selon un système de hiérarchie. Certaines personnes instruites produisent les liaisons sans enchaînements (bien que ce soit tout de même rare) et on trouve aussi des liaisons sans enchaînement avec *quand* et obligatoirement avec *vingt*. Les personnes peu instruites ou analphabètes font des erreurs avec les liaisons facultatives, erreurs similaires à celles de jeunes enfants d'âge pré-scolaire. Les erreurs sont prévisibles et peuvent s'expliquer par les effets de fréquence, d'analogie et de hiérarchie.

Les personnes moins éduquées (Laks, 1983) ont moins entendu certaines séquences avec des liaisons facultatives que les séquences avec des liaisons obligatoires, de la même

façon que les enfants de milieu défavorisé ont moins entendu de séquences et en produisent également moins que leurs pairs de milieu favorisé (Nardy, 2003; Chabanal, 2007). La fréquence de perception a des répercussions sur la fréquence de production.

L'étude des erreurs des adultes et des enfants nous renseigne sur le fonctionnement de la liaison et confirme ici notre hypothèse des liaisons formées selon le MBU. Celui-ci explique d'abord l'évolution de la langue et pourquoi certaines structures apparaissent, disparaissent ou fusionnent. Il explique comment les enfants apprennent une langue (Ellis, 2002 et Tomasello 2000a, 200b) et les liaisons n'y font pas exception. Les erreurs que font les enfants sont directement liées aux séquences qu'ils ont le plus entendues.

Le MBU permet d'expliquer les liaisons interdites ou les liaisons flottantes. Par exemple, on ne fait pas de liaison avec les mots d'origine étrangère et avec les noms propres qui commencent par une semi-consonne ou un h aspiré. Toutefois, si ces noms sont très fréquents, ils sont enregistrés dans le lexique des locuteurs comme appartenant à une autre catégorie et ils sont traités comme les mots fréquents. C'est ce qui explique la tendance à faire des liaisons avec les mots les plus fréquents ayant un h aspiré tels que *les haricots*, ou la tendance à dire *le h* (plutôt que *l'h*) ou *le hiatus* (plutôt que *l'hiatus*). C'est aussi la raison pour laquelle on peut faire la liaison avec des noms propres d'origine étrangère commençant par un h, si ceux-ci sont fréquents et connus de la majorité des francophones (*Henri, Henriette*).

Le MBU explique pourquoi la liaison avec *très* est obligatoire mais facultative avec *trop* ou d'autres adverbes et pourquoi elle varie selon les prépositions, les adverbes, les personnes et les temps. On utilise plus certaines personnes (3<sup>e</sup> personne) et certains temps (présent) que d'autres, et ce sont ces formes qui ont le plus de liaisons (Bybee, 2001). Il y a plus de répétitions en parole spontanée qu'en lecture et, dans cette parole spontanée, les contextes de liaisons possibles sont en plus faible pourcentage, pour plusieurs raisons.



Les groupes de mots les plus fréquemment employés ensemble fusionnent (*ch'uis*, pour *je suis*) ne permettant pas de liaisons après cette fusion. Cela nous renseigne sur le lien entre les mots. La négation ne fusionne jamais avec un mot, et n'a pas vraiment de lien avec des mots en particulier. Elle s'ajoute à des groupes déjà formés. C'est la raison pour laquelle son taux de cooccurrence est plus bas.

La séquence *est arrivée* a un taux de cooccurrence (d'après *Lexique*) de plus de 1000, alors *pas arrivé* a un taux de 405. Cela signifie que même si elle est relativement fréquente, elle l'est bien moins que d'autres concurrentes. Dans la hiérarchie des séquences les plus fréquemment trouvées ensemble, [pas + part. passé], n'est pas la plus élevée, parce que le participe passé est le plus employé avec l'auxiliaire (et non avec la négation).

Par contre, la séquence *pas encore* a un taux de cooccurrence de plus de 1000. Cela signifie que cette séquence a de plus fortes chances d'être présente dans le lexique des locuteurs, comme groupe de mots fréquemment trouvés ensemble. La liaison a donc beaucoup plus de chances d'être produite dans *pas encore*, que dans *pas arrivé*.

Tout ceci nous montre qu'on ne peut pas considérer un mot, ici l'adverbe *pas*, de façon isolée, mais qu'il faut considérer le mélange de plusieurs facteurs : la fréquence des mots, la fréquence de cooccurrence des mots, la fréquence des structures et des phonèmes. C'est l'ensemble de plusieurs facteurs qui permet aux locuteurs d'enraciner des structures et des séquences. On connaît déjà la loi de Zipf, comme l'indique Diessel (2007: 119) : « The correlation between frequency, length, and categorical status has become known as Zipf's (1935) law. ». Il existe en fait une loi qui regroupe encore plus de phénomènes : la loi du MBU qui regroupe la fréquence, l'analogie, la hiérarchie et la complexité, et toutes les composantes linguistiques sont étroitement liées.

C'est également la conclusion que tire Morin (2005 : 16) : « Cette circularité n'est pas qu'apparente : phonologie, morphologie, syntaxe et lexique ne peuvent être arbitrairement

dissociés les uns des autres dans l'analyse des facteurs responsables des différents types de liaison. »

Si la présence de certains facteurs est nécessaire à la production d'une séquence (ici une séquence avec liaison), l'absence d'autres facteurs (accent d'insistance, hésitation, pause, coup de glotte) l'est aussi. Le contexte d'énonciation, l'influence de l'écrit sur l'oral, la position des mots dans les phrases, la longueur des mots et des phrases et le nombre de liaisons avant ou après le mot liaisonnant entrent en considération. Il paraît difficile par exemple de faire la liaison avec des champs\_en jachère. Donc, ce qui rend la liaison difficile en empêche sa production, si la production crée des cas de cacophonie comme ceux mentionnés par Gadet (1989 : 56), tels que « vous êtes aux eaux [...] où l'on risque le calembour *zozo*, ou *au zoo* ».

Nous en venons à la conclusion que la production de liaisons dépendrait de la fréquence d'une séquence : fréquence du mot1, fréquence du mot2, fréquence de cooccurrence du mot1 et du mot2, fréquence et productivité d'une structure syntaxique, fréquence et productivité d'une structure phonétique, et de l'input langagier que le locuteur a reçu, en particulier en début d'acquisition, mais dépendrait aussi de l'absence d'éléments « perturbateurs » créant des difficultés d'articulation ou des cas de cacophonie.

D'après ce que nous avons vu dans ce chapitre, le locuteur utiliserait la banque de données de son lexique pour traiter simultanément des informations lexicales, phonétiques, syntaxiques, fréquentielles, stylistiques et sociolinguistiques et pour déterminer de façon inconsciente d'après sa mémoire auditive (et sa mémoire auditive et visuelle pour reconnaître les mots écrits), comment des séquences se prononcent.

Nous souhaitons analyser dans cette thèse le rôle des différents types de fréquence dans le traitement de la liaison des locuteurs francophones et surtout des locuteurs anglophones du français L2 pour vérifier les hypothèses amorcées dans ce chapitre.

## **CHAPITRE II**

### **L'ACQUISITION DES LIAISONS EN L1 ET EN L2**

## Introduction

Nous avons vu au chapitre 1 les principales théories relatives au MBU (Modèle Basé sur l'Usage). L'un des facteurs importants de ce modèle est le fait que les locuteurs traitent les informations d'après une base de données et d'après la fréquence des items et des structures permettant d'analyser des probabilités spécifiques à une langue. Les théories cognitives et psycholinguistiques sur l'acquisition du langage soutiennent que toutes les unités linguistiques proviennent de l'utilisation de la langue.

L'apprentissage d'une langue consiste à apprendre les représentations de ces associations qui reflètent les probabilités d'occurrences de mappages (*mapping*) entre la forme et la fonction. La fréquence est un facteur majeur dans l'acquisition parce que les règles d'une langue, à tous les niveaux d'analyse (de la phonologie à la syntaxe), sont des régularités des structures qui émergent de l'analyse de toute la vie de l'apprenant pour en déduire des caractéristiques distributionnelles de l'input langagier. La compréhension d'une langue est déterminée par l'énorme quantité d'informations statistiques relatives au comportement d'items lexicaux dans cette langue.

Pierrehumbert (2001) explique les différentes approches de la phonologie stochastique, mettant en relief le rôle de la fréquence dans l'acquisition, la reconnaissance et la production de divers domaines linguistiques :

The classical generative models are non-probabilistic. Any given sequence is either well-formed under a grammar, or it is completely impossible. Under this approach, statistical variation in observed data is viewed as related to variation in performance, rather than illuminating core competence. In contrast, work on sound structure in intellectual circles outside of generative linguistics proper has used probabilistic models for many decades. This line of research has established that the cognitive representation of sound structure is probabilistic, with frequencies playing a crucial role in the acquisition of phonological and phonetic competence, in speech production and perception, and in long-term mental representations. (Pierrehumbert, 2001 : 195)

Pierrehumbert (2001) souligne le rôle de la fréquence dans une interprétation probabiliste basée sur des exemplaires, impliquant le traitement d'un très grand nombre de

données non seulement pendant la prime enfance, lors du développement du langage, mais aussi tout au long d'une vie:

Establishing mental representations of phonetic distributions, including the contextual factors which play important roles, requires an immense amount of experience and considerable sophistication in encoding this experience. Although initial progress on acquiring these distributions is one of the earliest accomplishments of language acquisition - with considerable progress by 8 months (see review in Vihman, 1996) - adult mastery of allophony, stress / accent, and phonetic precision continues to develop for 6 to 12 year-olds [...] There is also evidence that updating of these probability distributions continues throughout adult life. (Pierrehumbert, 2001 : 196)

Pierrehumbert (2001 : 197) parle aussi d'effets probabilistes cumulatifs : « Of particular importance to the theoretical architecture is the existence of cumulative probabilistic effects (phenomena in which the probabilities associated with two different constraints combine to yield the likelihood of the outcome) ».

La fréquence des structures (*type frequency*) lui semble plus importante (p.197), que la fréquence des items (*token frequency*), tout comme Bybee (2001), car même si certains items sont très fréquents, ces quelques mots ne permettent pas une généralisation à d'autres items. S'il est difficile de prouver que la fréquence des structures est plus importante que la fréquence des items, ce qui est sûr, c'est que la fréquence des structures est un facteur très important.

Selon Pierrehumbert (2001 : 198), la productivité joue un rôle majeur : une structure doit avoir suffisamment d'items pour être productive. Plusieurs facteurs sont ainsi reliés : la fréquence des structures, leur productivité, mais aussi leur transparence sémantique et phonologique : « Frequency is known to play a role in the cognitive representation of morphophonological relationships. The acquisition of any given rule depends on having a sufficient number of examples (although it is important to note that other factors such as phonological and semantic transparency also play a role, with the result that frequency is not sufficient to predict order of acquisition). »

Les études en acquisition des liaisons chez les enfants L1 nous permettent de mieux comprendre le phénomène de la liaison et son traitement et les phénomènes d'acquisition en général. Observer les étapes par lesquelles les enfants passent et les erreurs qu'ils commettent nous renseigne sur les difficultés inhérentes aux liaisons elles-mêmes, ainsi que sur le processus neurolinguistique en jeu.

Nous verrons dans ce chapitre quelques études influentes en acquisition des liaisons chez les enfants L1 et analyserons les erreurs les plus courantes dans le cadre du MBU. Cela nous servira de base pour analyser ensuite les productions d'adultes L1 et d'apprenants L2.

## **1. L'acquisition des liaisons chez les enfants francophones**

### **1.1. Principes généraux d'acquisition d'après le MBU**

Comme nous l'avons vu dans le chapitre 1, les études en psychologie dans le cadre du MBU révèlent que l'apprentissage des enfants s'opère dans un processus global d'acquisition qui est commun à tous les enfants et à toutes les langues.

Vers 2-4 ans le jeune enfant apprend des items et des tronçons préfabriqués, sur une base lexicale. Il apprend des mots et des séquences au coup par coup. L'enfant les utilise sans vraiment en maîtriser les structures et le fonctionnement. Les items appris correspondent aux structures et aux items les plus fréquents dans son entourage. L'enfant apprend des groupes de mots comme s'ils ne formaient qu'un seul mot.

Pour qu'un item ait été acquis, il faut que l'enfant l'ait entendu un grand nombre de fois (nombre qui reste d'ailleurs à déterminer). Il apprend alors des items et des tronçons, sans être à ce stade capable de construire des règles et de les appliquer à des items jamais entendus. Il ne peut, à ce stade, que reproduire ce qui a été entendu. C'est d'ailleurs la raison pour laquelle, il fait peu d'erreurs à cette étape de son développement.

Puis, vers 4-5 ans, il met en place un système de catégories abstraites, dans lesquelles il va y insérer et classer les items existants et les nouveaux items au fur et à mesure de leur acquisition. Il formule alors des hypothèses d'après ce qu'il a entendu précédemment, et c'est à ce stade qu'on observe le plus d'erreurs. En effet, l'enfant tente de généraliser des règles à des items qu'il n'a jamais rencontrés dans une structure. Ou bien, au contraire, il formule des hypothèses quant à une structure et tente de la mettre en pratique alors que les règles sont erronées ou incomplètes et que des items sont irréguliers.

Pour que l'enfant soit capable de maîtriser une structure, il doit avoir acquis assez d'items pour construire une classe, mais aussi avoir entendu cette structure suffisamment de fois pour pouvoir y adjoindre les items entendus. De plus, pour qu'il ait acquis une structure, il faut que celle-ci soit saillante, que l'enfant l'ait remarquée comme ayant du sens pour lui, et y ait prêté attention. L'acquisition d'items et de structures s'opère simultanément bien que l'enfant apprenne d'abord des items et une fois un grand nombre d'items appris, il maîtrise la structure.

La productivité et la complexité d'une structure jouent un rôle majeur dans l'acquisition. Par exemple, si une structure 1 regroupe plus d'items qu'une structure 2, la structure 1 rentre en compétition avec la structure 2 qui a moins d'items. L'enfant va d'abord apprendre la structure 1 qui regroupe plus d'items, surgénéraliser cette règle, et l'utiliser là où il faudrait utiliser la structure 2.

Par exemple, la structure [dét. pluriel + nom commençant par voyelle] est plus fréquente que la structure [dét. pluriel + nom commençant par un h aspiré]. Ainsi les enfants apprendront d'abord *les amis*, *les avions*, avec un phonème de liaison /z/ qu'ils généraliseront au h aspiré pour dire comme on le trouve dans les exemples de Pallaud et Savelli (2001 : 124-127) : « des-z-haricots (9 ans) [...] trois-z-hamsters (âge non précisé) [...] deux-z-héros (11 ans) [...] »

On retrouve le même phénomène dans l'acquisition de la morphosyntaxe et du lexique. Par exemple, dans l'acquisition de l'auxiliaire aux temps composés, puisque l'auxiliaire *avoir* est utilisé avec beaucoup plus de verbes que ne l'est l'auxiliaire *être*, la structure [aux. avoir + p. passé] est plus productive que la structure [auxiliaire être + p. passé]. Ainsi, les enfants utilisent l'auxiliaire *avoir* au lieu de l'auxiliaire *être* dans les exemples de Pallaud et Savelli (2001 : 129) : « j'ai allé jouer... je m'ai trompé ».

De même, les verbes du 1<sup>er</sup> groupe sont plus nombreux que les verbes des autres groupes, donc les enfants généralisent leurs conjugaisons aux verbes des 2<sup>e</sup> et 3<sup>e</sup> groupes. De plus, les formes de la troisième personne sont plus utilisées que celles des autres personnes (ce qu'on a vu dans le corpus PFC et dans l'étude de Bybee, 2001). Pour les verbes irréguliers, à l'oral, les trois premières personnes ont généralement la même forme phonétique. Par exemple, les conjugaisons des trois personnes du singulier du verbe savoir : *je sais, tu sais, il sait* se prononcent toutes les trois /se/. Les enfants apprennent donc d'abord la 3<sup>e</sup> personne du singulier qu'ils généralisent ensuite à toutes les personnes. De même le présent étant plus fréquent que les autres verbes, surtout dans l'environnement langagier des enfants, ils généralisent les règles du présent aux autres temps.

En effet, en analysant les erreurs enfantines, Pallaud et Savelli (2001) remarquent que les enfants ont tendance à :

- uniformiser les verbes du 2<sup>e</sup> et 3<sup>e</sup> groupe en verbe du 1<sup>er</sup> groupe : « suivre » devient « suiver » ;
- à conjuguer la 1<sup>e</sup> personne comme la 3<sup>e</sup> personne : « je l'a vu quatre fois [...] quand j'aura fait des enfants [...], je pourra [...]
- à régulariser des verbes irréguliers : « vous disez » ;
- à faire des fautes de conjugaison avec le passé simple avec la même tendance de régularisation consistant à traiter les verbes irréguliers comme des verbes réguliers et des conjugaisons du 2<sup>e</sup> et 3<sup>e</sup> groupe comme celles du premier : « il coura [...], Mickey disa [...], il faisa [...], sa trompe se défaisa [...], l'éléphant sorta, [...] mes parents connaissèrent [...]. (Pallaud et Savelli : 2001 : 130-131)

Ces erreurs reflètent l'input des enfants ainsi que le processus général d'apprentissage commun à tous les aspects linguistiques, dont la tendance à simplifier les règles et à utiliser, par analogie, les structures les plus fréquentes, même pour les structures les moins fréquentes.



Si l'enfant a entendu suffisamment de fois des items et des structures, il maîtrise cette structure, bien qu'il puisse encore faire des erreurs, mais de façon exceptionnelle (seulement pour des items rares ou des exceptions). Il ne fait plus celles qu'il faisait auparavant. Grâce à la fréquence des structures et des items entendus, il précise et renforce les nouvelles classes créées et en crée de nouvelles. Les structures et les items les plus fréquents sont les premiers à être utilisés et acquis. Les enfants sont alors capables d'utiliser les nouveaux items (ce qu'on voit dans des exercices de pseudo-mots) qu'ils viennent d'entendre, et de généraliser des règles pour créer des tronçons qu'ils n'ont jamais entendus.

L'interaction joue un rôle déterminant pour « valider » les structures. Non seulement cela renforce les acquis, mais cela permet aussi de rectifier les erreurs grâce au renforcement positif ou négatif de l'entourage. Ainsi l'enfant abandonne les structures fautives et ne garde dans son inventaire que les structures grammaticalement correctes des natifs.

En fait, il est possible, d'après une base de données reproduisant l'input que reçoivent les enfants d'un certain âge, de prédire les productions et les erreurs des enfants relatives à des décisions lexicales. Dufau et al. (2006) ont mis au point un logiciel qui génère un comportement de décision lexicale qui recrée l'apprentissage des formes orthographiques des apprentis lecteurs français. Les résultats sont similaires aux performances des enfants de 6 à 11 ans.

Cette étude a mis en évidence les effets de la fréquence lexicale : les mots de haute fréquence sont reconnus plus aisément que ceux de basse fréquence. Les mots qui apparaissent le plus fréquemment dans une langue sont perçus plus rapidement et produisent moins d'erreurs de reconnaissance que des mots qui apparaissent moins fréquemment. L'effet du voisinage orthographique est moins visible. Les résultats renforcent l'hypothèse selon laquelle la construction du lexique orthographique chez les enfants est un processus guidé principalement par les données auxquelles ils sont exposés.

## 1.2. Lacunes dans le développement langagier de l'enfant

Pour que le processus d'acquisition d'une langue s'opère « normalement », il faut que certaines conditions soient présentes : un input riche et varié, des séquences bien formées, une variété de locuteurs et de situations d'énonciation, etc. Il faut aussi que certaines conditions soient absentes. En effet, il ne faut pas qu'il y ait de bruits (au sens d'entraves), car ceux-ci peuvent empêcher le système de se mettre en place (Barlow et Kemmer, 2000).

Si un enfant a des difficultés d'audition, d'attention, d'apprentissage, de comportement, s'il n'a pas l'input langagier nécessaire, s'il n'entend pas assez de séquences bien formées et s'il ne reçoit pas le feedback nécessaire, cela va avoir un impact sur son apprentissage. Si les difficultés sont isolées ou momentanées, elles peuvent être corrigées par des stratégies de compensation, mais si les difficultés subsistent, l'apprentissage est ralenti ou même interrompu. C'est notamment le cas pour les enfants sourds ou autistes.

En cas de contact présent, mais moins riche avec des locuteurs natifs, l'apprentissage est plus lent. Nous verrons plus loin que les enfants dont les parents sont de milieu social plus bas ont des performances langagières similaires à celles des enfants plus jeunes (Hoff, 2002).

Si l'enfant manque d'interaction avec des locuteurs natifs (pour le rôle de l'interaction, voir par exemple Long, 1996), le développement est très fortement ralenti. On voit cela par exemple dans les recherches de Sachs, Bard et Johnson (1981) sur des enfants qui ne sont pas sourds mais dont les parents le sont. Selon cette étude, une exposition passive à la langue, telle que l'écoute de la télévision, ne suffit pas à apprendre et maîtriser une langue, ce qui met en évidence le rôle de l'interaction. Si l'enfant a moins d'une dizaine d'années, il aura du retard dans son acquisition, mais il sera encore possible de compenser ses lacunes si des mesures sont prises pour l'aider dans son développement en répondant à ses besoins spécifiques. Si l'enfant a passé cette période critique, les lacunes seront irrémédiables et l'enfant aura des lacunes dans son apprentissage.

L'enfant traite les informations à partir d'un système de probabilités grâce à une banque de données, dans laquelle les informations sont organisées dans un système de classe et de hiérarchie selon leur fréquence. Les bébés sont déjà capables de traiter des banques de données (voir par exemple, Saffran, Aslin et Newport, 1996) pour la reconnaissance de syllabes ou d'autres tâches de reconnaissance. Mais si l'enfant n'a aucun input langagier et ne se trouve pas dans des conditions « normales » d'apprentissage, son cerveau ne pourra pas se constituer une banque de données, créer des classes et des catégories abstraites et finalement il ne pourra pas se développer normalement (Itard, 1962 ; Curtiss, 1977 ; Rymer 1993).

Dans des conditions de maltraitance extrême ou un enfant n'a aucun input linguistique ou social, l'enfant ne peut pas apprendre à parler, marcher, manger, etc., ce qu'on voit avec Victor, l'enfant sauvage de l'Aveyron ayant vécu isolé dans la forêt (Itard, 1962) ou Génie (Curtiss, 1977 ; Rymer, 1993), la petite fille qui vivait toute seule dans une pièce sans aucun contact avec le monde extérieur, attachée à une chaise. Ses parents lui apportaient des bouillies sans lui parler, et si la petite fille tentait de produire des sons, elle était sévèrement frappée.

En cas de privation extrême, le cerveau des enfants ne peut se développer normalement. Victor et Génie ne sont jamais arrivés à produire autre chose que quelques mots isolés prononcés avec peine et de façon à peine compréhensible. Si le cerveau ne reçoit pas l'input nécessaire, celui-ci ne peut créer une banque de données, acquérir des items et des structures, et développer des hypothèses et des catégories. C'est ainsi que les linguistes et psychologues ont observé une période critique en L1 après laquelle il n'est plus possible d'acquérir une langue, période se situant environ entre 10 et 12 ans et selon certains, même avant pour l'acquisition de la phonétique. La fréquence et la qualité de l'input déterminent les productions des locuteurs en L1.

### 1.3. Recherches en acquisition chez les enfants L1

On retrouve des phénomènes similaires dans l'acquisition du lexique, de la morphosyntaxe et de la phonétique. Monnin, Loevenbruck et Beckman (2007) ont analysé le rôle de la fréquence dans l'acquisition des phonèmes chez les enfants. Ils ont basé leur étude sur trois corpus : 1) LEXIQUE (<http://www.lexique.org>), corpus de langue adulte basé sur 3200 textes écrits, contenant 31 millions d'items ; 2) LEXIQUE 3, corpus de langue adulte orale contenant 14.7 millions d'items, conçu à partir de sous-titres de films reflétant la langue orale dans des dialogues ; 3) un corpus conçu à partir d'enregistrements d'adultes (le plus souvent la mère parlant à l'enfant) représentant l'environnement langagier de l'enfant.

Monnin, Loevenbruck et Beckman (2007) concluent que lorsque certains phonèmes sont moins fréquents dans une langue, ils sont acquis plus tard et entraînent plus d'erreurs, ce qui confirme le rôle de la fréquence dans l'acquisition de certains aspects phonétiques. Les auteurs ont étudié les différences qui surviennent dans l'acquisition de langues différentes. Puisque la fréquence de certaines consonnes est différente selon les langues, ils ont ainsi pu étudier le rôle de la fréquence des phonèmes présents dans la langue ambiante dans l'acquisition de phonèmes de langues différentes.

Monnin, Loevenbruck et Beckman (2007 : 4) observent qu'en début d'acquisition, la complexité des phonèmes joue un rôle primordial dans le développement phonétique. En effet, les phonèmes et séquences les plus « difficiles phonétiquement » sont acquis beaucoup plus tard que les autres, quel que soit l'environnement langagier de l'enfant. Puis, une fois les phonèmes demandant le moins d'efforts articulatoires appris, et ceux-ci sont les mêmes à travers les langues, le rôle de la fréquence se fait ressentir dans l'acquisition des phonèmes « plus difficiles » : « For example, sibilant fricatives and affricates are mastered later than stops [1, 2], and dorsal stops are mastered later when the following vowel is /i/ as compared to /u/ [3]. Second, some of these “phonetically difficult” sounds and sound sequences are

acquired relatively sooner in some languages than in others. For example, /v/ seems to be mastered earlier in Swedish, Estonian, and Bulgarian than it is in English [4]. »

Dans les trois corpus, /s/ et /k/ sont plus fréquents que /t/ en français. L'étude de la production de séquences CV en début de mots produites par les enfants français montre que les phonèmes /k/ et /s/ sont en majorité correctement produits (sauf cas particuliers). De plus, le phonème /k/ (dans les contextes /a/, /i/, /y/, /u/) entraîne moins d'erreurs que les phonèmes /t/ et /s/. /k/ n'est jamais remplacé par /t/, alors que /t/ est souvent remplacé par /k/. Cela explique le taux élevé de productions correctes de ces deux phonèmes.

Le taux élevé de phonèmes correctement réalisés se remarque d'autant plus que /s/ et /k/ sont considérés comme étant « difficiles » d'un point de vue articulatoire. On remarque le même phénomène en grec et en japonais, langues pour lesquelles /k/ est plus fréquent que /t/. Par contre le taux de productions correctes et le taux d'erreurs diffèrent de l'anglais, langue dans laquelle /k/ est moins fréquent que /t/.

Il ressort donc que l'environnement langagier joue un rôle majeur dans l'ordre d'acquisition des phonèmes, de la précision et justesse (*accuracy*) des productions, dans les erreurs et les substitutions, ce que l'on peut observer d'après la fréquence des phonèmes de chaque langue.

Un phénomène similaire se présente avec l'acquisition des liaisons. Hahmad (1993) a remarqué que dans 20 heures de parole adulte, les contextes syntaxiques de liaisons obligatoires (dét + nom, pronom + verbe, etc.) étaient plus fréquents que les contextes facultatifs et qu'en situation informelle, seules 10% des liaisons facultatives étaient réalisées.

Adda-Decker, Boula de Mareüil et Lamel (1999), et Boula de Mareüil et Adda-Decker (2002) ont relevé que la liaison est plus fréquente avec les mots les plus fréquents (75%), qu'avec les mots rares (20%) et que les liaisons sont toujours réalisées avec les déterminants.

Toutefois, même si ces résultats indiquent que la fréquence jouerait un rôle important

dans la production de liaison, il semble que ce ne serait pas le seul, sinon on obtiendrait 100% et non 75%. C'est pourquoi nous avons souhaité développer plus en détail le rôle de la fréquence dans cette thèse, pour déterminer avec plus de précision les facteurs participant à la production de liaisons.

Booij et de Jong (1987) remarquent que les liaisons sont réalisées à 100% dans quatre contextes : après déterminant, après pronom préverbal, entre verbe et pronom et dans certaines expressions figées. Durand et Lyche (2008) confirment l'étude de Booij et de Jong et en viennent à la conclusion que les liaisons dites « obligatoires » sont réalisées à plus de 90% (sauf en cas de contexte interdisant la liaison, tel que h aspiré, semi-consonne, etc.) et sont en accord avec Delattre (1951), excepté avec les prépositions monosyllabiques, les adjectifs antéposés et la construction impersonnelle *c'est*. Nous avons vu dans le chapitre précédent que le contexte [adj. + nom] est un contexte plus rare. Quant à la fréquence de réalisation des liaisons avec les prépositions monosyllabiques, celle-ci dépend de chacune des prépositions, puisque certaines prépositions sont beaucoup plus fréquentes que d'autres.

Selon *Lexique*, la fréquence (dans les films) de *en* est 5689, la fréquence de *dans* 4658, de *chez* 842, *depuis* 483 et *pendant* 312. Alors que la liaison est toujours réalisée avec *en* et *dans*, elle l'est rarement avec *depuis* et *pendant* et elle varie avec *chez* (pas de liaison entre *chez* et un nom propre).

Selon le MBU, les enfants devraient apprendre les liaisons d'après un ordre déterminé par la fréquence. Les liaisons obligatoires se trouvant dans les structures les plus fréquentes : [dét. + nom], avec clitique, etc., devraient être apprises avant les liaisons facultatives. Certaines liaisons facultatives sont plus fréquentes que d'autres. Par exemple, [*c'est* + groupe nominal (GN)] est une structure plus fréquente et à l'intérieur de cette séquence, le mot1 et le mot2 ont une plus forte cohésion et un taux de cooccurrence plus élevé que la structure [nom pluriel + adj.].

D'après le MBU, les enfants devraient donc apprendre les structures simples et fréquentes, puis les structures simples et un peu moins fréquentes et en dernier les structures complexes et plus rares. C'est précisément ce qui se passe. Les enfants francophones apprennent d'abord les liaisons les plus fréquentes : les liaisons obligatoires. Et à l'intérieur de cette catégorie, ils apprennent d'abord le contexte le plus fréquent qui est le contexte [article + nom]. Chevrot (2002), Chevrot et Fayol (2000) et Chabanal (2003) ont remarqué que les enfants apprennent les liaisons obligatoires avant les liaisons facultatives.

Selon Wauquier-Gravelines et Braud (2005 : 63), les enfants se basent sur un modèle morphologique et « une capacité très précoce à établir des généralités phonologiques sur la base d'un input phonétiquement riche mais phonologiquement lacunaire ». Nous ne partageons pas l'opinion que l'input est « phonologiquement lacunaire », nous avons discuté au chapitre 1 de la richesse de l'input. Wauquier-Gravelines et Braud (2005 : 57) observent que 64% des productions enfantines sont des productions de substantifs. Selon les auteures, les enfants utilisent plus de noms, donc créent d'abord des théories abstraites sur la syllabation et la formation de mots en utilisant des groupes nominaux, donc des contextes [dét. + nom].

Même si nous ne partageons pas l'avis des auteures quant aux explications (principalement morphologiques) de l'ordre d'acquisition des enfants, les éléments mentionnés dans leur étude n'en sont pas moins révélateurs : les enfants apprennent d'abord les noms et donc les articles qui vont avec les noms. Nous sommes par contre d'accord que les consonnes rajoutées au mot<sup>2</sup> sont comme si elles faisaient partie de l'article, comme un « proto-déterminant » (selon les termes de Wauquier-Gravelines et Braud, 2005), ce que Chevrot et Fayol (2000), Chabanal (2003) et Dugua (2005) considèrent comme des allomorphes du mot<sup>2</sup> que l'enfant utilise selon les situations.

Morin (2005) a remarqué que les enfants passent par des étapes et font des erreurs telles que « gros t-enfant » et « soudain t-intérêt ». Celui-ci a aussi mentionné que les adultes font des erreurs similaires de type « quatre z-enfants, comme z-arbres, des chefs d'état z-africains, des symptômes gastro-z-intestinaux » et que les personnes peu éduquées font peu de liaisons en particulier dans le contexte [adj. + nom]. Il conclut aussi qu'une solution phonologique ou morpho-phonologique (d'épenthèse) n'explique pas ces phénomènes.

Dugua (2005) a analysé avec des tâches de segmentation, d'apostrophe et de liaison le processus de segmentation de 200 enfants âgés de 1 à 6 ans. Les résultats de son étude suggèrent que les enfants passent par plusieurs étapes dans leur développement. D'abord ils répètent des mots et séquences de mots, puis construisent progressivement des structures abstraites basées d'abord sur des mots, puis sur des séquences, et enfin sur des structures à emplacements ouverts à certaines catégories et enfin des « schémas composites ». Dugua (2005) observe que pour les enfants :

[...] il s'agit de choisir entre deux constructions en compétition : faire la liaison ou ne pas la faire. La liaison présente une double particularité qui la distingue des autres phénomènes de sandhi. Elle implique, d'une part, une disjonction entre frontière lexicale et syllabique, et d'autre part, la présence d'un segment variable à cette frontière. Cette double particularité complexifie la tâche de l'enfant, en témoignent les deux principaux types d'erreurs repérées depuis longtemps (Grégoire, 1947 ; Chevrot, 2002) : les erreurs par remplacement (/dɔnurs/ pour /dɔzurs/) et les erreurs par omission (/dɔurs/ prononcé sans /n/). [...] (Dugua, 2005 : 2-3)

Dugua (2005) relève trois étapes : 2-3 ans, 3-4 ans, et 5-6 ans avec « une mise en place progressive » entre 2 et 5 ans :

Dans la tâche d'apostrophe, les enfants à 2-3 ans connaissent le mot « isolé » (sans phonème de liaison), mais semblent considérer le mot à initiale vocalique comme une variante parmi les autres variantes ayant la consonne de liaison adjointe au mot. A 5 ans, les liaisons obligatoires « sont quasiment acquises (80.30%) mais les liaisons facultatives sont peu réalisées jusqu'à 4 ans. Puis à partir de cet âge les performances progressent régulièrement jusqu'à 6 ans. (Dugua, 2005 : 6)

Les résultats de Dugua (2005 : 8) sur les tâches d'apostrophe (de type « appelle l'animal : ours, viens ici »), de segmentation et de liaison avec dénomination d'image (de



type « qu'est-ce qu'il y a sur l'image ? Deux ours ») proposées à un échantillon de 200 enfants âgés de 2;4 à 6;1 ans font apparaître les étapes d'acquisition par lesquelles ils passent. D'une part, ils apprennent d'abord les liaisons les plus fréquentes. D'autre part, ils passent par des étapes qui sont en accord avec le MBU et les recherches de Tomasello (2003). Enfin, les erreurs et productions des enfants sont représentatives de leur stade de développement et de leur input langagier :

[...] l'enfant produirait des séquences liées (Mot1-CL-Mot2) sans erreur de type substitution ou omission, puisque les premières acquisitions langagières se feraient sur la base de séquences perçues dans l'input et mémorisées telles quelles dans le lexique, de façon non analysée. [...] Les enfants [...] font des erreurs impliquant des CL, erreurs qui suggèrent qu'une analyse s'opère sur ces séquences. Commenceraient alors les processus de segmentation [...] des séquences liées entendues dans l'input [...] Ainsi, l'enfant mémorise et stocke dans son lexique différentes variantes à consonne initiale d'un même mot, variantes issues des différents contextes dans lesquels elles ont été entendues. [...] Cette étape se mettrait en place essentiellement grâce au contact avec les séquences bien formées rencontrées dans l'input. (Dugua, 2005 : 8)

Dugua (2006) a analysé les erreurs de 512 enfants âgés de 2 à 6 ans et a corrélié l'effet de la fréquence à leurs erreurs. Pour déterminer la fréquence dans l'input des enfants, les auteurs de l'étude ont demandé aux personnes proches des enfants (parents), de déterminer, d'après leur intuition, la fréquence de certaines séquences utilisées dans l'entourage des enfants. Cela a permis d'établir les séquences les plus souvent entendues par les enfants. Les enfants devaient décrire des images dans les contextes : [un + Mot2] et [deux + Mot2].

Le but de l'étude était d'étudier les erreurs par remplacement du type *le navion* ou *un zavion*. Les résultats obtenus ont montré que les erreurs produites étaient fortement et significativement corrélées avec la fréquence des formes entendues. Si dans leur entourage, ils avaient le plus souvent entendu « un n-avion », ils faisaient des erreurs du type « les n-avions » et s'ils avaient entendu le plus souvent « des z-avions » ils faisaient des erreurs de type « un z-avion ». Cette étude est ainsi en accord avec les prédictions du MBU, puisqu'elle indique que les erreurs sont directement corrélées à la fréquence des séquences entendues.

La fréquence en réception et en production joue un rôle majeur dans l'acquisition de séquences chez l'enfant (ici les liaisons dans certaines séquences). Cette fréquence est aussi reliée à l'âge d'acquisition. Selon Ferrand, Grainger et New (2003), les enfants apprennent en premier les mots les plus fréquents, les plus courts et ceux avec lesquels ils sont familiers. Toutefois, ils apprennent aussi des mots plus rares (d'après les critères généraux de fréquence sur *Lexique*), tels que *fée*, *sorcière*, *géant* qui sont des mots importants pour les enfants car plus saillants. Grainger et New (2003) en concluent que si l'on sépare les facteurs de fréquence et d'âge d'acquisition, on remarque que l'effet de l'âge d'acquisition est plus important pour les mots plus rares. Par contre, pour les mots fréquents, ce facteur est moins important :

Pour la lecture à voix haute, il semble que les mots appris en premier sont traités différemment des mots appris plus tard, ce que l'on voit dans les tâches de reconnaissance lexicale et dans le traitement en lecture à haute voix. [...] les effets de l'âge d'acquisition et de la fréquence sont additifs, ils affectent des niveaux différents de traitement, le premier affectant la production phonologique des mots, le second affectant la reconnaissance visuelle des mots. [...] L'effet de fréquence affecterait soit le temps d'accès à la représentation orthographique lexicale du mot, ou bien les connexions entre la reconnaissance visuelle et la phonologie. Autrement dit, l'âge d'acquisition affecterait la « sortie » du système de lecture, tandis que la fréquence affecterait « l'entrée » du système de lecture. (Grainger et New, 2003 : 453-454)

En outre, il semble que l'âge d'acquisition dépende de toute façon de l'input langagier de l'enfant puisque celui-ci ne peut pas apprendre des mots s'il ne les a pas entendus auparavant. Ainsi, on peut difficilement séparer le rôle de la fréquence du rôle de l'âge d'acquisition. Par contre, on peut dire que la fréquence et l'âge d'acquisition dépendent tous deux de l'environnement langagier qui est lui-même déterminé par les parents de l'enfant et leur milieu socioéconomique.

D'après le MBU, plusieurs facteurs rentrent en jeu : non seulement la fréquence relative des mots, mais aussi la saillance, la complexité des mots et structures, etc. Si bien que si un mot est « plus rare » tel que *sorcière*, dans la conversation journalière des adultes (sur *Lexique*), d'une part, cela ne veut pas dire qu'il soit plus rare dans l'environnement de

l'enfant. En effet, *Lexique* représente pour la plupart, la parole adulte. *Lexique* n'est donc pas forcément représentatif de l'input langagier de l'enfant.

D'autre part, quand bien même les mots seraient effectivement moins fréquents dans l'environnement langagier des enfants, des mots tels que *sorcière, magicien, prince, fée*, etc. sont pour les enfants plus « saillants ». Même si ces mots sont plus rares, leur saillance compense leur fréquence relative et favorise donc leur acquisition.

De plus, il est possible que, si un enfant a entendu ces mots lorsque quelqu'un lui lisait une histoire, dans un dessin animé ou dans un film, la mémoire visuelle associée aux émotions de l'enfant l'ait aidé à retenir ces mots moins fréquents, mais dont le sens était plus évident et important que des mots plus fréquents, mais plus insignifiants ou plus abstraits (malheureux, vivant...). Ainsi, la fréquence serait associée à d'autres facteurs tels que la saillance.

Enfin, les parents et le milieu socioéconomique de la famille jouent un rôle dans l'input langagier des enfants, car s'ils sont exposés à plus de livres, histoires, dessins animés, chansons, discussions, etc., ce que suggère Hoff (2002), alors les enfants de milieu favorisé devraient apprendre ces mots avant ceux de milieu défavorisé. Si les parents lisent plus d'histoires aux enfants, ces derniers entendront plus souvent certains mots tels que « sorcière, géant, etc. »; si les enfants regardent des émissions et dessins animés pour enfants, ils entendront plus souvent ces mêmes mots qui sont présents dans la plupart de ce que les parents leur ont lu. Les dessins animés pourraient renforcer ce qui a été entendu dans les livres, expliquer certains concepts abstraits, etc. et *vice versa*. Plus les activités sont nombreuses, plus elles peuvent renforcer, améliorer et accélérer l'acquisition en améliorant la fréquence de l'input et en le rendant aussi plus saillant et pertinent pour l'enfant.

Il semble donc, comme le mentionnent Ferrand, Grainger et New (2003), qu'il est difficile de séparer le facteur fréquence de mot (et cooccurrence de mot) et le facteur âge d'acquisition. Toutefois ce qui est sûr, c'est que l'input langagier joue un rôle dans la

fréquence des mots entendus et produits et que les enfants de milieu favorisé ont des connaissances et compétences lexicales plus développées que leurs pairs au même âge. Puisque les enfants de milieu favorisé apprennent des mots et les structures allant avec ces mots avant leur pairs, ils bénéficieront à la fois de la fréquence et de l'âge d'acquisition.

Dans le développement langagier, l'acquisition de la morphosyntaxe, de la prononciation des mots et de la lecture est liée à l'acquisition du lexique. Il faut d'abord connaître les mots pour en connaître les propriétés morphosyntaxiques, leurs relations avec d'autres mots (telles que les liaisons), leur prononciation, et pour être capable de les reconnaître à l'oral et à l'écrit. Il faut connaître les mots à l'oral pour les reconnaître à l'écrit, d'où la difficulté à lire des pseudo-mots. Il faut d'abord avoir entendu les formes orales avant de les reconnaître à l'écrit (voir la médiation phonologique dont discutent Sprenger-Charolles, et al., 2003). Sinon, comment un locuteur saurait-il par exemple que le d de *quand* ou de *grand* se prononce /t/ en liaison ? En outre, les enfants L1 apprennent à parler avant d'apprendre à lire et à écrire.

D'un autre côté, les connaissances de l'écrit permettent aux enfants de savoir comment les mots s'écrivent, de faire le lien entre l'oral et l'écrit et d'établir ainsi des représentations abstraites des structures plus rares, notamment avec les liaisons facultatives et interdites, même si la majorité est acquise par l'oreille, indépendamment de l'écriture (puisque les enfants produisent déjà toutes les liaisons obligatoires à 6 ans).

L'apprentissage de la lecture dépend, entre autres, de la capacité à lire et à décoder les mots et les séquences de mots. Sprenger-Charolles, et al. (2003) ont suivi le développement d'une cinquantaine d'enfants de la 1<sup>e</sup> à la 4<sup>e</sup> année (CP au CM1 en France). D'après leurs recherches, ce qui prédit les futures compétences en lecture d'un apprenti lecteur, c'est sa capacité à décoder les sons (*phonological processing*) en début d'apprentissage et plus tard.

Les enfants se servent de la correspondance graphie/phonie pour deviner comment s'écrivent les mots, et ce, même lorsqu'ils reconnaissent les mots familiers, avec des pseudo-mots ou des mots irréguliers dont la graphie ne correspond pas à la prononciation, tels que *femme* ou *album*. La lecture à voix haute aide à l'écriture plus que la lecture à voix basse. Cela permet à l'enfant de faire le rapport entre graphie et phonie, alors que s'il lit à voix basse il reconnaît seulement la graphie.

Sprenger-Charolles, Siegel et Bonnet (1998) ont observé les stratégies qu'utilisent les enfants en première année d'école en France (appelé en France, le Cours Préparatoire ou CP). Pour tester certaines hypothèses, les chercheurs ont demandé aux enfants de lire des mots dont les graphèmes correspondaient aux phonèmes (des mots réguliers), des mots dont les graphèmes ne correspondaient pas aux phonèmes (des mots irréguliers tels que *femme*), des mots contenant plusieurs lettres pour un phonème, par exemple les lettres *ou* pour /u/ ou *ch* pour /ʃ/, des pseudo-mots parmi lesquels des mots analogues à d'autres mots connus et des mots non-analogues. Parmi les mots « réels » réguliers et irréguliers, il y avait une catégorie de mots à haute fréquence et une catégorie de mots à basse fréquence.

Les auteurs ont ainsi pu tester le rôle de l'opacité des graphèmes, de l'analogie, de la fréquence et de « la médiation phonologique » chez des enfants de CP entre janvier et juin. Ils ont comparé les stratégies utilisées et celles non utilisées. Au début (en janvier), les élèves utilisent surtout la médiation phonologique, c'est-à-dire qu'ils ont recours aux sons pour savoir comment écrire et ils ont besoin de connaître le rapport phonème/graphème.

Puis, en juin, on voit l'effet de la fréquence des mots rencontrés pendant ces quelques mois ainsi que le recours à l'analogie. Les enfants ont quand même recours à la médiation phonologique. Le rapport phonème/graphème les aide à se constituer un lexique orthographique qui les aide ensuite, par analogie, à lire et écrire des mots nouveaux. Plus leur lexique orthographique est grand, plus ils peuvent utiliser l'analogie et comparer les mots

nouveaux aux mots familiers de ce lexique, pour les reconnaître et les prononcer correctement.

En milieu de CP, les scores en lecture de mots réguliers et de pseudo-mots ne diffèrent pas et sont supérieurs à ceux relevés pour les mots irréguliers, alors très faibles. Puis, en fin de CP, les mots réguliers sont alors mieux lus que les pseudo-mots, eux-mêmes mieux lus que les mots irréguliers. Les auteurs de l'étude expliquent que c'est probablement le cas parce que les mots réguliers bénéficient à la fois de la régularité et de la fréquence d'exposition. Ni la fréquence d'exposition, ni la régularité ne suffisent à elles seules à améliorer les compétences, comme l'indique la progression plus faible relevée pour les mots irréguliers d'une part, et pour les pseudo-mots d'autre part.

D'après cette étude, la capacité de décodage détermine le succès de l'apprentissage de la lecture. Les meilleurs décodeurs précoces étant ceux qui progressent le plus, y compris en lecture de mots irréguliers et en compréhension écrite. Ces différents résultats permettent de comprendre pourquoi l'enseignement systématique et précoce des correspondances grapho-phonémiques est ce qui aide le plus efficacement les élèves dans l'apprentissage de la lecture.

Ceci a plusieurs implications pour notre étude sur les liaisons. La lecture dépend des capacités de décodage des sons et de la médiation phonologique, c'est-à-dire de la capacité à faire le lien entre graphème et phonème. Nous avons donc ici plusieurs types de fréquences : celle des graphèmes et celle des phonèmes. Pour pouvoir reconnaître et produire des séquences correctes, les enfants doivent les avoir entendues un grand nombre de fois. Pour pouvoir lire rapidement et correctement, les enfants doivent connaître le lien graphème-phonème, être capables de reconnaître les mots écrits et donc de les avoir vus et lus souvent, et de faire le lien avec d'autres graphèmes, par analogie. Nous voyons donc que plusieurs sortes de fréquences jouent un rôle dans la capacité de produire des séquences correctes avec des liaisons en parole spontanée et en lecture.

Chevrot, Dugua et Fayol (2005 : 41) ont étudié l'acquisition de la liaison chez les enfants et ont remarqué, en se référant à Dugua (2002), que les items les plus fréquents sont mémorisés en premier. Leur étude avec des enfants d'âges divers a montré que le pourcentage d'erreurs sur les liaisons diminue au fur et à mesure que l'enfant grandit, entre 2-3 ans et 5-6 ans, comme on le voit dans le tableau 2.1.

âge	% d'erreurs
2-3 ans	35.4%
3-4 ans	19%,
4-5 ans	8%
5-6 ans	2.8%

**Tableau 2. 1. % d'erreurs de liaisons chez les enfants L1. Chevrot, Dugua et Fayol (2005 : 41)**

De plus, les enfants semblent apprendre d'abord les mots avec leur liaison : « *nours viens ici* » et ensuite le mot séparément (vers 6 ans). De même, Chevrot, Dugua et Fayol (2005 : 43), en se référant à Dugua (2002), relèvent que l'acquisition des liaisons est plus rapide après un déterminant, avec les liaisons obligatoires et les liaisons les plus fréquentes.

Chevrot, Dugua et Fayol (2005 : 43) en concluent que « les liaisons obligatoires progressent entre 2 et 5 ans, alors que l'évolution des facultatives débute à 3-4 ans et se poursuit jusqu'à 6 ans », comme on peut le voir dans le tableau 2.3. Il apparaît ici que l'acquisition des liaisons, tout comme les erreurs de liaisons sont directement liées à l'input reçu et au rôle de la fréquence. Les liaisons les plus fréquentes (obligatoires) sont acquises en premier et les enfants font moins d'erreurs avec les liaisons obligatoires. Les liaisons facultatives sont moins fréquentes que les liaisons obligatoires, elles sont apprises plus tard que les liaisons obligatoires et provoquent un plus grand nombre d'erreurs de substitution (cuirs et velours).

% de liaisons obligatoires justes réalisées	% de liaisons facultatives justes réalisées
2-3 ans 36%	2-3 ans 12%
3-4 ans 54%	3-4 ans 14%
4-5 ans 80%	4-5 ans 22%
5-6 ans 83%	5-6 ans 33%

**Tableau 2. 2. % de liaisons obligatoires et facultatives des enfants L1. Chevrot, Dugua et Fayol (2005 : 43)**

Si le contexte langagier dans lequel baigne l'enfant joue un rôle sur l'acquisition et la production linguistique, comment peut-on évaluer l'input de l'enfant ?

Tout d'abord, Hoff (2002) a comparé les conditions langagières d'enfants issus de différentes classes sociales. Les parents à statut social plus élevé (parents cadres) fournissent à leurs enfants un environnement langagier plus riche : le temps de parole est plus élevé, les contextes de discussions sont plus variés, ils sollicitent davantage la participation active des enfants dans la conversation, par exemple en leur posant des questions, et utilisent eux-mêmes une langue plus riche, plus variée et de niveau plus élevé que les parents de niveau socio-économique plus bas.

Chevrot, Dugua et Fayol (2005 : 7), en se référant à Nardy (2003), ont étudié le rôle de l'input langagier de 189 enfants de 2 à 6 ans de classe sociale différente (tableau 2.3.).

L'input langagier joue un rôle déterminant dans les productions justes de liaisons obligatoires et il varie selon la classe sociale.

Les enfants qui ont des parents cadres auraient (d'après Hoff, 2002) entendu davantage de parole, entendu et participé à plus de discussions avec les parents, et produisent un taux de liaisons nettement plus élevé que les enfants de parents issus de la classe ouvrière. Cette différence se fait ressentir très tôt (2-3 ans) et elle est toujours visible à l'âge de 5-6 ans.

	Parents cadres	parents ouvriers
2-3 ans	67 %	25 %
3-4 ans	74 %	60 %
4-5 ans	83 %	65 %
5-6 ans	96 %	86 %

**Tableau 2.3. Production de liaisons obligatoires d'enfants de 2-3 ans à 5-6 ans selon la classe sociale des parents. Chevrot, Dugua et Fayol (2005 : 7)**

On peut ainsi en déduire que l'acquisition de la liaison dépend de l'environnement langagier de l'enfant, de la qualité et des productions présentes dans l'input, de la fréquence d'écoute et de la fréquence de production de l'enfant. Si les différences sont énormes à 2-3 ans, elles s'estompent un peu lorsque l'enfant va à l'école. Ceci s'explique en partie parce que



l'input langagier durant la journée devient plus uniforme à l'école. Toutefois des différences sont toujours présentes. Par exemple, les jeunes adolescents de banlieues défavorisées (Laks, 1983) ne produisent que 4% de liaisons facultatives, ce qui représente un pourcentage nettement plus bas que les résultats obtenus dans toutes les autres études de corpus francophone.

De plus, on ne peut pas penser que ces différences s'estomperont complètement pour plusieurs raisons. Les différences socioculturelles persisteront et même s'accroîtront au fil des années, selon que les enfants continuent leurs études, ou arrêtent l'école très tôt.

Chabanal (2007) a analysé l'influence du milieu et de l'héritage linguistique dans la production phonologique de jeunes enfants et en est venu aux mêmes conclusions que Chevrot, Dugua et al. (2005) et Hoff (2002). Il a d'abord étudié le rôle de l'input dans la production de liquides /l/ et /R/ en finale complexe (*cartable, fenêtre...*) chez deux enfants de classes sociales différentes, âgés de 40 mois, pendant 10 mois, avec un enregistrement tous les deux mois. Le but était de comparer les productions des enfants de 3 ans à leur input langagier (leurs parents).

Il ressort de l'étude que, dès l'âge de 4 ans, des différences notables se remarquent dans la production de liquides en finales complexes. L'enfant au statut social le plus élevé a réalisé beaucoup plus de liquides que l'autre enfant. Il n'a jamais omis le groupe de consonnes finales, et l'autre enfant l'a omis dans 44% des cas (*cartab* ou *carta* au lieu de *cartable*). Cela suggère qu'il a moins entendu les formes complètes et ne les a pas assimilées dans son lexique. Chabanal (2007 : 4) en conclut que « Ces éléments nous permettent de souligner que l'héritage linguistique apparaît bien dès les trois premières années de l'enfant. Cela signifie donc que la langue de l'école peut entrer en conflit avec la langue parlée par l'enfant dans son milieu. »

Chabanal (2007) a ensuite étudié le rôle de l'input dans la production de liaisons facultatives de deux groupes de 12 enfants de milieu favorisé et 12 enfants de milieu défavorisé âgés de 8 ans. En répondant aux questions des évaluateurs, les enfants se sont trouvés dans différentes situations langagières : S1 : dialogue informel, S2 : exercices structuraux et S3 : évaluation : juste ou faux (*match guise*).

Comme on le voit dans le tableau 2.4., les enfants de milieu favorisé produisent plus de liaisons facultatives que ceux de milieu défavorisé et les différences sont d'autant plus remarquables dans la tâche de reconnaissance, où les enfants de milieu favorisé reconnaissent 66% de liaisons facultatives comme justes, alors que les enfants de milieu défavorisé n'en reconnaissent que 13%.

Liaisons facultatives	S1	S2	S3
Pourcentages de LF réalisées justes (milieu favorisé)	22% (94/434)	16.5% (18/108)	66% (71/108)
Pourcentages de LF réalisées justes (milieu défavorisé)	15% (58/396)	7.5% (8/108)	13% (14/108)

**Tableau 2.4. Pourcentages de liaisons facultatives réalisées justes selon les milieux. Chabanal (2007 : 6)**

À partir de ces résultats nous pensons que les enfants de milieu favorisé ont entendu des liaisons facultatives plus souvent et dans des contextes plus variés, c'est pourquoi ils les reconnaissent quand ils les entendent. Cela semble aussi indiquer (notre interprétation) qu'à tous les niveaux, la réception précède la production.

Cela expliquerait les différences entre les tâches 1, 2 et 3 et expliquerait aussi les différences entre les enfants de milieu favorisé et ceux de milieu défavorisé. Puisqu'ils n'ont pas entendu ces liaisons dans leur entourage, ils ne peuvent ni les reconnaître comme correctes puisqu'elles ne font pas partie de leur lexique, ni les produire.

Chabanal (2007 : 7) conclut de ces deux études que : « Ces difficultés proviennent de la faiblesse du bain linguistique (pauvre et peu varié), d'une conscience stylistique moins accrue et de la transmission précoce d'un vernaculaire. La reconnaissance de ce dernier

comme langue seconde proprement dite (Verdelhan, 2002) devrait amener à une réflexion plus efficace. Notamment sur la variété stylistique du français oral et ses ‘sanctions sociales’.»

Chabanal et Embarki (2002) ont cherché à savoir si les enfants de 9 ans utilisent plus de liaisons facultatives dans des situations plus formelles, et si les enfants de milieu favorisé produisent plus de liaisons facultatives que les enfants de milieu défavorisé. Pour cela, les enfants ont participé à quatre activités les conduisant à quatre situations : Situation 1: dialogue informel ; Situation 2: exercices structuraux (demande plus formelle) ; Situation 3: évaluation de la forme juste ; Situation 4: récitation d'une poésie en classe.

Les résultats de l'étude (tableau 2.5.) suggèrent, d'une part, que les enfants des deux milieux sont capables d'adapter leur langue à une situation langagière en utilisant plus de liaisons dans une situation plus formelle qui requiert plus de liaisons facultatives, (récitation de poésie) ; d'autre part, que les enfants de milieu favorisé produisent plus de liaisons facultatives que les enfants de milieu défavorisé et que ces derniers font plus d'erreurs de liaisons que les enfants de milieu favorisé.

Situations	S1 - dialogue	S2 - ex. structuraux	S3- évaluation	S4 -poésie
Milieu favorisé	33,3%	16,6%	83,3%	100%
Milieu défavorisé	41,6%	50%	75%	91,6%
Ensemble de l'échantillon	37.4%	33.3%	79.1%	95.8%

**Tableau 2.5. Liaison avec le verbe « être ». Chabanal et Embarki (2002 : 178)**

Nous pouvons faire quelques observations à partir des résultats de ce tableau. Rappelons quelques faits vus au chapitre précédent concernant la liaison avec *être* chez les adultes. D'après les règles de Delattre (1951) et de l'Académie Française, la liaison est obligatoire avec *est*. D'après Ågren (1973) cette liaison est une liaison facultative fréquente puisque la liaison avec [est + participe passé] se fait dans 98.6% des cas et dans 97% des cas pour *est* en général. Durand et Lyche (2008 : 46), d'après les données du corpus PFC, concluent que la liaison avec *est* n'est réalisée que dans 30% des cas. La liaison avec *être*

varie beaucoup, selon de nombreux critères, notamment le registre de langue, le type d'énonciation, le niveau d'étude des locuteurs, etc.

Les situations langagières des études en question pourraient correspondre aux situations de l'étude de Chabanal et Embarki (2002) qui pouvaient inciter les enfants à produire plus de liaisons avec *être*. Comme la liaison avec *être* est l'une des plus fréquentes parmi les liaisons facultatives, les enfants de milieu défavorisé ont peut-être remarqué une différence entre l'input de l'école (l'institutrice, les autres élèves), et l'input à la maison. Ils en ont peut-être déduit (inconsciemment) que dans une situation particulière (avec des étrangers, à l'école, etc.), il fallait faire la liaison avec *être*.

Si l'on considère les situations 1 (dialogue) et 2 (exercices structuraux) qui ont une moyenne respective de 37.4% et de 33.3% de liaisons réalisées, cela indiquerait que les productions des enfants seraient supérieures à celles des locuteurs du projet PFC. Il convient donc de considérer ces résultats avec précaution. Il est peu probable que des enfants de 9 ans fassent plus de liaisons facultatives que des adultes, et ce, quel que soit leur niveau socioéconomique.

Nous avons vu au chapitre précédent que les résultats dépendent de facteurs tels que la personne et le temps. Ainsi, la 1<sup>e</sup> personne du singulier et les temps du passé ou du futur engendreraient trois fois moins de liaisons que les 3<sup>es</sup> personnes et le temps présent. Il est possible que les jeunes locuteurs de milieu défavorisé aient acquis dans leur lexique des séquences spécifiques à certaines situations (de l'école), sans qu'elles soient suffisamment présentes dans leur environnement familial, du type : Qu'est-ce que *c'est* (enseignant) ? *C'est un* ... (élève). Bien sûr à la maison, l'entourage utilise l'inévitable verbe *être*, mais sans doute avec beaucoup moins de liaisons. Si bien que les pourcentages de liaisons réalisées avec *c'est* ne sont, à notre avis, pas forcément représentatifs de la maîtrise de liaisons facultatives fréquentes, mais seulement représentatives de la liaison dans certains cas avec le verbe *être*.

Par contre, nous pouvons déduire de ces études que les jeunes locuteurs de milieu favorisé peuvent mieux adapter leur discours à une situation langagière et à un registre correspondant et semblent mieux maîtriser la variation sociolinguistique. De plus, les situations les plus représentatives sont celles pour lesquelles il y a un véritable enjeu, situations similaires à des situations scolaires, situations d'examen, etc. Autrement dit, ce qui est représentatif des différences entre les groupes, c'est surtout la plus grande variation du groupe de milieu favorisé.

Chabanal et Embarki (2002 : 179) mentionnent aussi les liaisons fautives, même s'ils ne donnent pas de précisions, et donnent des exemples d'erreurs présentes dans les productions des enfants de milieu défavorisé : « il prend (z) un ballon, je suis (t) allé... », alors que de telles erreurs sont absentes des productions des enfants de milieu favorisé.

Les résultats de l'étude suggèrent d'une part une prise de conscience de tous les locuteurs d'un marqueur sociolinguistique qui les incite à produire plus de liaisons dans des situations plus formelles, mais aussi un manque de familiarité des enfants de milieu défavorisé avec ces structures, ce qui les pousse à produire des liaisons dont ils ne maîtrisent pas les structures et produire ainsi des liaisons fautives. Comme le font remarquer Chabanal et Embarki (2002 : 179) : « Si le jeune locuteur a donc coutume de ne pas prononcer les variables normatives coûteuses sur le plan articulatoire, parce qu'il ne les entend pas chez lui, la difficulté sera plus grande et les taux d'erreurs également. Les erreurs présentes ici en termes de liaisons fautives sont à noter justement chez les locuteurs de milieux défavorisés. »

Nous pouvons ainsi conclure que plusieurs éléments peuvent représenter les connaissances des locuteurs : la production de certains marqueurs, tels que les liaisons facultatives dans un certain contexte, les erreurs typiques d'un contexte, d'une situation ou d'un milieu social, et enfin, la capacité à adapter leur discours à une situation particulière ou,

comme le dit Laks (2000 : 3) : « La capacité à adapter son style linguistique au niveau exigé par l'organisation sociale de la situation de parole. »

Dugua et Chabanal (2006) ont évalué les productions de liaisons de jeunes enfants dans deux études. Les deux études comprenaient une tâche d'apostrophe, avec 7 figurines d'animaux dont 4 avec voyelle initiale (âne, écureuil, éléphant, ours) et 3 à consonne initiale (cochon, chien, perroquet). Le but était de déterminer si les enfants possèdent dans leur lexique les mots (mot2) à initiale vocalique sans « préfixe » et, s'ils utilisent un phonème de liaison, quel est-il et pourquoi ils l'utilisent plutôt qu'un autre.

Pour l'étude transversale, 178 enfants âgés de 2;6 à 6;1 ans ont été testés. Les résultats de l'étude (tableau 2.6.) indiquent que les enfants ont dans leur lexique la forme à initiale vocalique puisque environ le tiers des enfants produisent le mot isolé. Toutefois, ils ont appris les mots avec leur consonne de liaison, puisque 39.2% d'entre eux utilisent le mot2 avec une consonne de liaison. Les enfants utilisent de plus en plus les mots à initiale vocalique, au fur et à mesure qu'ils grandissent. Les jeunes enfants utilisent plutôt les mots à initiale consonantique et en particulier avec le phonème /n/. 10 enfants ont utilisé le phonème /z/ et 7 ont utilisé le phonème /t/.

	forme à initiale vocalique (ours !)	forme à initiale consonantique (nours !)	forme avec déterminant (un ours ! l'ours !).
<b>2;6-3;1</b>	32.0% (30.2)	39.2% (30.2)	28.8% (32.3)
<b>3;6-4;3</b>	56.8% (40.6)	18.6% (25.9)	24.6% (34.4)
<b>4;4-4;9</b>	68.0% (36.9)	13.3% (19.3)	18.8% (29.8)
<b>4;10-5;5</b>	73.3% (34.2)	12.5% (22.6)	14.2% (28.5)
<b>5;6-6;1</b>	81.2% (27.1)	3.8% (11.1)	15.0% (24.2)

**Tableau 2.6. Résultats de l'étude transversale (moyennes et écarts-types). Dugua et Chabanal (2006 : 2)**

Pour l'étude longitudinale, 20 enfants âgés de 2;10 à 5;9 ans ont été suivis pendant 4 ans, lors de cinq observations. Là aussi (tableau 2.7.), les enfants utilisent surtout les variantes à initiale consonantique et en particulier la consonne /n/: 102 productions en /n/ sur 130

productions à consonne initiale. Les enfants utilisent de moins en moins de mots à initiale consonantique au profit de mots à initiale vocalique (81.8% vers 6 ans).

	forme à initiale vocalique (ours !)	forme à initiale consonantique (nours !)	forme avec déterminant (un ours ! l'ours !).
<b>2;6-3;1</b>	50.0% (33.5)	28.6% (30.9)	28.2% (38.1)
<b>3;6-4;3</b>	47.8% (36.3)	28.6% (32.1)	23.6% (28.7)
<b>4;4-4;9</b>	58.3% (37.4)	19.6% (28.5)	22.1% (26.7)
<b>4;10-5;5</b>	84.4% (20.6)	7.8% (16.3)	6.6% (14.1)
<b>5;6-6;3</b>	81.8% (22.7)	0% (0)	17.6% (22.7)

**Tableau 2.7. Résultats de l'étude longitudinale (moyennes et écarts-types). Dugua et Chabanal (2006 : 3)**

L'étude longitudinale suggère qu'en début d'apprentissage les enfants ont tendance à segmenter la consonne de liaison avec le mot<sup>2</sup>. Mais à partir de 6 ans, c'est-à-dire à partir du moment où ils apprennent à lire et à écrire, ils n'attachent plus du tout un phonème de liaison au mot<sup>2</sup> (0%). On remarque également un apprentissage en cloche, où entre les âges de 3 et 4 ans, les enfants produisent moins de variantes à initiale vocalique et plus de variantes à initiale consonantique. Ceci est représentatif de l'apprentissage des enfants qui au début utilisent seulement les éléments entendus, puis commencent à formuler des hypothèses et font à ce stade plus d'erreurs en général.

Cette étude met en relief le rôle de l'environnement langagier dans l'apprentissage de l'enfant. Les enfants entendent dans leur entourage les mots à initiale vocalique principalement avec un déterminant, et avec un phonème de liaison. Ils l'entendent beaucoup plus rarement, si ce n'est dans l'étude expérimentale, sous forme « isolée », sans article et sans phonème de liaison. C'est la raison pour laquelle ils apprennent d'abord les mots avec le déterminant et la consonne de liaison l'accompagnant.

D'après les résultats de Mallet (2008 : 211), la liaison la plus réalisée par les adultes francophones est la liaison avec /n/ qui est réalisée à 46.4% (par rapport à 38.7% pour /t/ et 14.7% pour /z/). Il est fort probable que l'enfant entende encore plus souvent dans son environnement langagier des formes au singulier. Il semble plus probable que l'enfant entende

plus souvent « l'âne » ou « un âne » que « les ânes », « l'éléphant » ou « un éléphant » plus souvent que « les éléphants ». Il suffit pour cela de se référer aux histoires pour enfants (Winnie l'ourson par exemple).

La principale structure qui serait le plus à même de mettre en relief un nom à initiale vocalique sans phonème de liaison serait la structure [adj. + nom], comme par exemple « un joli arbre », ou avec un autre phonème d'enchaînement tel que « bel arbre, ou nouvel arbre ». Mais, comme Mallet (2008) et Durand et Lyche (2008) l'ont remarqué, cette structure est rare. Enfin, si les adjectifs sont au pluriel, ils permettent la liaison tout comme avec un article. Les noms se trouvent donc principalement avec des déterminants, et avec la structure *c'est*.

L'apprentissage de la liaison semble se développer selon la fréquence des contextes et des items rencontrés. Plus l'enfant rencontre de contextes et d'items, plus il peut compléter son inventaire et créer des catégories abstraites. Dugua et Chabanal (2006) en concluent :

Nous pouvons remarquer des analogies avec la théorie probabiliste des exemplaires. Kirchner (1999) postule que les représentations lexicales sont complexes et qu'il existerait différents exemplaires d'un même mot stockés lors des différentes rencontres avec celui-ci. En outre, pour Pierrehumbert (2001) le choix de l'exemplaire dépend de sa récence et de sa fréquence dans le lexique mental. Comme l'écrit Dugua (2002) "l'influence de la fréquence revient à postuler que plus un exemplaire est souvent activé, plus il sera fort". Les données obtenues dans cette recherche suggèrent que des formes différentes de Mots 2 coexistent dans le lexique des jeunes enfants. Nos résultats sont par conséquent compatibles avec le fonctionnement lexical impliquant la mise en œuvre de connexions entre items. (Dugua et Chabanal, 2006 : 4).

Chevrot, Dugua et Fayol (2005) ont analysé les données des études précédentes concernant l'acquisition des liaisons chez les enfants de 2-6 ans et leurs erreurs et en déduisent que c'est l'exposition aux séquences mot1-mot2 bien formées rencontrées dans l'environnement langagier qui conditionne l'apprentissage de la liaison chez l'enfant. En ce sens, ils vont à l'encontre d'une conception morphologique telle que celle de Wauquier-Gravelines (2002, 2003). Cette hypothèse de la fréquence de l'input est en accord avec la théorie des exemplaires décrite par Pierrehumbert (2002).



Chevrot, Dugua et Fayol (2005) ont comparé les résultats de Dugua (2002) à certaines informations présentes dans le dictionnaire de Boë et Tubach (1992) qui contient les fréquences d'occurrence de mots. Il s'avère que les progressions des enfants correspondent à la fréquence des mots. Plus les mots sont fréquents, plus les enfants apprennent les liaisons correspondantes tôt :

[...] *deux* apparaît 65 fois suivi de la liaison /z/ réalisée, alors que *petit* n'apparaît que 8 fois devant la liaison /t/ réalisée. [...] L'acquisition est plus précoce pour la liaison suivant *deux* [...]. A l'inverse, les progrès sur la liaison suivant *petit* ne sont pas significatifs [...] Il s'ensuit que les liaisons justes sont plus fréquentes pour *deux* que pour *petit*, [...] Si les jeunes enfants maîtrisent la liaison /z/ de *deux* plus précocement que la liaison /t/ de *petit*, c'est parce qu'ils rencontrent souvent *deux* suivi d'une CL dans des séquences mot1-mot2. (Chevrot, Dugua et Fayol, 2005 : 8)

Chevrot, Dugua et Fayol (2005 : 8) ont également remarqué que les enfants maîtrisent les liaisons obligatoires vers l'âge de 6 ans. Ils progressent régulièrement dans leurs productions de liaisons obligatoires justes et font de moins en moins d'erreurs jusqu'à l'âge de 6 ans : « Les études d'enregistrements familiaux montrent que la réalisation des liaisons obligatoires [...] atteint au plus tard vers 6 ans le score de 100 % observé chez l'adulte. Dans cinq études de cas transversales, (Chevrot, 2002), les pourcentages individuels de liaisons justes sont de 84 % vers 3 ans 6 mois, de 98 % vers 5 ans 9 mois, de 99 % vers 6 ans 11 mois. »

Chevrot, Chabanal et Dugua (2007) ont étudié la production de liaisons obligatoires et facultatives ainsi que les erreurs de 5 enfants âgés de 3 à 11 ans dans une étude transversale. De cette étude (tableau 2.8.) ressortent plusieurs faits marquants. Les enfants maîtrisent les liaisons obligatoires à 6-7 ans. Le nombre de séquences justes avec liaisons est proportionnel à l'âge. Par exemple, à 6-7 ans, ils produisent 100% de liaisons obligatoires avec 11 expressions figées; à 10-11 ans, ils produisent le même pourcentage (100%) de liaisons justes, mais cette fois ils produisent 45 expressions figées justes, soit quatre fois plus. Soit ils produisent les expressions figées justes, soit ils ne les produisent pas du tout, ils n'en

produisent pas avec des erreurs. Autrement dit, on voit ici, que ce n'est pas seulement les pourcentages qui rentrent en jeu, mais aussi le nombre de séquences justes.

Sujet et âge approximatif	A. (3–4 ans)	B. (5–6 ans)	C. (5–6 ans)	D. (6–7ans)	E. (10–11ans)
<b>Liaisons obligatoires justes</b>					
a. après un déterminant ( <i>les_ ours</i> )	<b>77%</b> 119/154	<b>91%</b> 61/67	<b>97%</b> 71/73	<b>97%</b> 71/73	<b>100%</b> 100/100
b. après un pronom clitique	<b>91%</b> 131/144	<b>100%</b> 140/140	<b>100%</b> 205/205	<b>100%</b> 205/205	<b>99%</b> 263/264
c. dans une expression figée ( <i>tout_ à coup</i> )	– 1/1	<b>100%</b> 13/13	– –	<b>100%</b> 11/11	<b>100%</b> 45/45
d. entre verbe et clitique post-verbal ( <i>prenez_en ! vient_il ?</i> )	- -	1/1 -	- -	- -	- -
<b>Total liaisons obligatoires justes</b>	<b>84%</b> <b>251/299</b>	<b>97%</b> <b>215/221</b>	<b>99%</b> <b>187/189</b>	<b>99%</b> <b>287/289</b>	<b>99%</b> <b>408/409</b>
<b>Liaisons facultatives justes</b>					
e. après un adjectif pré-nominal ( <i>un petit_ ours</i> )	<b>48%</b> 11/23	– 3/4	– 6/8	– 4/7	– 3/10
f. après un nom pluriel ( <i>des enfants_ idiots</i> )	<b>2%</b> 1/58	– 2/11	<b>18%</b> 2/8	<b>0%</b> 0/13	<b>0%</b> 0/17
g. après une forme du verbe avoir ( <i>ils ont_un. . .</i> )	– 0/7	<b>0%</b> 0/15	– 0/7	<b>0%</b> 0/50	<b>0%</b> 0/28
h. après une forme du verbe être ( <i>ils sont_ ici</i> )	<b>0%</b> 0/49	<b>5%</b> 4/76	<b>0%</b> 0/50	<b>12%</b> 10/84	<b>20%</b> 39/198
i. après une forme d'un autre verbe ( <i>il vient_ aussi</i> )	<b>1%</b> 1/94	<b>2%</b> 2/106	<b>1%</b> 1/78	<b>0%</b> 0/118	<b>5%</b> 10/214
j. après mot invariable ( <i>en_Espagne, quand_un. . .</i> )	<b>3%</b> 3/99	<b>7%</b> 8/121	<b>26%</b> 19/73	<b>24%</b> 35/146	<b>12%</b> 27/220
<b>Total liaisons facultatives justes</b>	<b>5%</b> <b>16/330</b>	<b>6%</b> <b>19/333</b>	<b>12%</b> <b>28/224</b>	<b>12%</b> <b>49/418</b>	<b>12%</b> <b>79/687</b>
<b>Erreurs de substitution en contexte de liaison obligatoire</b> ( <i>un ours produit un –z-ours où /z/ remplace le /n/attendu</i> )	<b>5%</b> 16/299	<b>0%</b> 0/221	<b>0%</b> 0/189	<b>0%</b> 0/289	<b>0%</b> 0/409
<b>Erreurs de substitution en contexte de liaison facultative</b> ( <i>je suis allé produit je suis –t-allé</i> Où /t/ remplace le /z/attendu)	<b>3%</b> 9/330	<b>0.6%</b> 2/333	<b>0%</b> 0/224	<b>0%</b> 0/418	<b>1.5%</b> 10/687
<b>Erreurs d'omission en contexte de liaison obligatoire</b> ( <i>des ours prononcé [deurs] sans liaison</i> )	<b>11%</b> 32/299	<b>3%</b> 6/221	<b>1%</b> 2/189	<b>1%</b> 4/289	<b>0.2%</b> 1/409

Tableau 2.8. Cinq études de cas : liaisons correctes et erreurs. Chevrot, Chabanal et Dugua (2007 : 112)

En ce qui concerne les liaisons facultatives, les enfants en produisent peu, même à 10-11 ans. Il est surprenant qu'après les mots invariables dont les prépositions monosyllabiques (qui sont souvent considérées comme obligatoires), ils n'en produisent, à 10-11 ans, que 12%, ce qui est moins que les enfants plus jeunes (26% à 5-6 ans). Toutefois, nous pouvons considérer que parmi les invariables, toutes les prépositions et conjonctions ne lient pas de la même manière. Il est possible que les enfants plus jeunes aient utilisé quelques items les plus fréquents engendrant un plus grand nombre de liaisons (*dans un, en un*), alors que les enfants plus âgés auraient utilisé une plus grande variété d'invariables, dont des invariables polysyllabiques. Il semble que les enfants n'aient pas terminé leur apprentissage des liaisons facultatives à 10-11 ans, puisque le taux de liaisons réalisées reste relativement faible.

Enfin, l'étude des erreurs fournit des informations sur le nombre de liaisons obligatoires et facultatives, et indique qu'entre 3-4 ans les enfants utilisent la moitié moins (330) de liaisons facultatives que ce que les enfants utilisent à l'âge de 10-11 ans (687). S'ils utilisent plus de liaisons facultatives à 10-11 ans qu'auparavant, il est possible qu'ils utilisent dans ce nombre, des structures moins fréquentes, causant ainsi un peu plus d'erreurs.

Lorsqu'ils sont plus jeunes, il est possible qu'ils utilisent les mêmes structures qui leur sont relativement familières, et fassent moins d'erreurs. Ainsi, il faut considérer le nombre (et le pourcentage) d'erreurs, mais aussi, le nombre et la complexité des structures utilisées. Un nombre restreint d'erreurs peut être signe d'une utilisation limitée de structures et une augmentation des erreurs peut être un signe d'une maîtrise d'une plus grande variété de structures. C'est pourquoi le nombre de structures (bien) utilisées peut être aussi important que le pourcentage de réalisations correctes.

La deuxième étude (tableau 2.9.) est une étude longitudinale de deux jeunes garçons de milieux sociaux différents entre 40 et 50 mois. Les parents de G. sont de classe moyenne (les deux parents ont bac +2), les parents de F. sont ouvriers et ont arrêté l'école à 14 ans.

L'étude longitudinale montre que, pour les deux enfants, la période 48-80 mois marque un seuil au-delà duquel les liaisons obligatoires sont acquises pour G, et pratiquement acquises pour F. Le taux d'erreurs diminue pour les liaisons obligatoires atteignant pratiquement 0% pour G. mais étant encore de 9% pour F.

Les enfants commencent leur acquisition des liaisons facultatives dans la tranche 1. F n'a pas encore atteint la tranche 2 qui est le niveau que G avait dans la tranche 1. Cela suggère que les enfants suivent le même développement, mais ceux de niveau social plus bas ont un développement plus lent. Ayant entendu moins souvent les séquences, ils ont besoin de plus de temps pour les avoir entendues le même nombre de fois. Un phénomène similaire s'observe pour les erreurs avec les liaisons obligatoires. F. fait approximativement le même nombre d'erreurs dans la tranche 2 que G. dans la tranche 1. En fait, les enfants de niveau social plus bas sont approximativement de même niveau que les enfants de statut social plus élevé, plus jeunes. Cela signifie qu'à âge égal dans une situation scolaire, les enfants de niveau social plus bas seront nettement désavantagés.

		<b>Tranche 1 40–47 mois</b>	<b>Tranche 2 48–50 mois</b>
<b>Liaisons obligatoires justes</b>	Enfant F.	<b>76%</b> 130/170%	<b>88%</b> 152/172
	Enfant G.	<b>89%</b> 145/162	<b>99%</b> 171/172
<b>Liaisons facultatives justes</b>	Enfant F.	<b>5%</b> 4/83	<b>6%</b> 4/66
	Enfant G.	<b>7%</b> 9/138	<b>21%</b> 22/104
<b>Erreurs de substitution en contexte de liaison obligatoire</b> (un ours produit un -z-ours où /z/ remplace le /n/attendu)	Enfant F.	<b>17%</b> 29/170	<b>9%</b> 16/172
	Enfant G.	<b>8%</b> 13/162	<b>0.5%</b> 1/172
<b>Erreurs de substitution en contexte de liaison facultative</b> (je suis allé produit je suis -t-allé , /t/ remplace le /z/attendu)	Enfant F.	<b>22%</b> 18/83	<b>27%</b> 18/66
	Enfant G.	<b>20%</b> 28/138	<b>28%</b> 29/104
<b>Erreurs d'omission en contexte de liaison obligatoire</b> (des ours prononcé [deurs] sans liaison)	Enfant F.	<b>6%</b> 11/170	<b>2%</b> 4/172
	Enfant G.	<b>0%</b> 0/172	<b>2%</b> 4/162

Tableau 2.9. Évolution longitudinale des liaisons obligatoires et facultatives et des erreurs chez deux enfants entre 40 et 50 mois. Chevrot, Chabanal et Dugua (2007 : 117)

Les résultats de l'étude longitudinale confirment ceux de l'étude transversale, et confirment elles-mêmes les études précédentes. A l'âge de 6 ans, la plupart des enfants ont acquis les liaisons obligatoires et font peu d'erreurs de liaison, mais cela dépend aussi du niveau d'étude des parents. Les productions dépendent de l'input reçu et les erreurs reflètent les tendances à utiliser les structures les plus fréquentes (donc plus accessibles) à la place de structures moins fréquentes.

Alors que l'acquisition des différentes structures de liaisons obligatoires est plutôt uniforme (presque 100% pour toutes les catégories (déterminant + nom, après un pronom clitique, après une expression figée...), on remarque, par contre, de la variation pour l'acquisition des liaisons facultatives. Par exemple, les enfants de 10-11 ans produisent 20% de liaisons facultatives justes après une forme du verbe *être* et 12% après une forme invariable, alors qu'ils n'en produisent aucune après un adjectif antéposé.

Comme les liaisons facultatives sont moins produites par les parents moins éduqués, ce qu'on voit dans les productions d'adultes en général, comme dans Ashby (1981, 1988, 2003), l'écart se creuse dans la production langagière d'enfants issus de différentes classes sociales. Les écarts que l'on trouve entre les différents types de liaisons facultatives sont proportionnels à la fréquence des structures, ce qui renforce le modèle de Bybee (2001) selon lequel l'enfant mémorise les séquences selon leur fréquence.

Chevrot, Chabanal et Dugua (2007) ont calculé le nombre de productions justes et les ont corrélés aux nombres de productions justes et au nombre d'items produits pour chaque structure. Ils ont pu établir un lien direct entre le nombre d'items produits pour chaque structure, et le nombre de liaisons justes. Cela signifie que, plus un enfant produit une structure, plus il la maîtrise, et ceci est aussi représentatif de son input langagier. Plus il l'entend, plus il la produit, plus il la produit, plus il la maîtrise :

Le classement des mots<sup>1</sup> selon les pourcentages de liaisons justes (colonne 2) est quasi identique à leur classement selon le rapport de la colonne 5, à l'exception du possessif *ses*. Plus un clitique ou un déterminant est employé dans des contextes de liaisons mot1-mot2 fréquents, plus la liaison est réalisée juste par l'enfant de 3 ans. [...] Ainsi, comme le prédit le modèle exemplariste et basé sur l'usage, un enfant de 3 ans réussit mieux les liaisons obligatoires dans des séquences lexicales mot1-mot2 qu'il produit souvent et dont on peut penser qu'elles sont souvent entendues. (Chevrot, Chabanal et Dugua, 2007 : 114-115)

La troisième partie de l'étude était l'analyse d'erreurs chez l'enfant S entre 2 ans 1 mois et 6 ans 4 mois, soit un total de 898 erreurs desquelles 437 erreurs ont été retenues, apparaissant toutes devant un nom masculin (les autres erreurs n'ont pas été analysées notamment à cause des enchaînements avec le féminin vs la liaison avec le masculin). Les erreurs ont été recueillies par le père de la fillette au fur et à mesure qu'il entendait sa fille les prononcer. Il y avait aussi d'autres restrictions qui ont permis de sélectionner 13 noms masculins apparaissant dans au moins 6 erreurs.

Les auteurs ont testé le lien entre les erreurs et l'input langagier de l'enfant. Pour déterminer les formes les plus susceptibles d'être entendues par l'enfant, les expérimentateurs ont demandé à des étudiants en sciences du langage (qui ne connaissaient pas le contenu de l'étude) de juger intuitivement si certains mots leur semblent être plus souvent utilisés au singulier ou au pluriel. De plus, ils ont cherché dans *Frantext*, les occurrences de ces mots au singulier et au pluriel, et ont ainsi déterminé quel contexte était le plus fréquent. Par exemple, le mot *ami* apparaît 3499 fois avec l'indéfini singulier *un*, 2254 fois avec le défini *le*, 2422 fois avec *des* et 1503 fois avec *les*. Puisque le mot *ami* est utilisé le plus souvent dans la séquence *un ami*, cela prédit que, s'il y a une erreur avec le mot *ami*, cela devrait être d'abord avec l'ajout d'un n (séquence la plus entendue, donc la plus accessible).

Les erreurs enfantines correspondent aux intuitions des sujets, qui correspondent elles-mêmes aux statistiques de fréquence de *Frantext*, quant à la fréquence des séquences entendues. Lorsque les sujets ont jugé qu'un mot est plus souvent utilisé au pluriel, les erreurs des enfants consistaient le plus souvent à remplacer la consonne de liaison par un /z/ au lieu

d'un /n/ ou si le mot est plus souvent rencontré au singulier, les enfants ont utilisé plus souvent la consonne de liaison /n/ au lieu de /z/.

Au début, l'enfant répète ce qu'il a entendu, mais les séquences possibles sont incomplètes dans la mesure où certaines n'ont pas été entendues assez souvent. Par exemple si la forme du singulier est moins fréquente que celle du pluriel, alors la forme la plus fréquente (ici la forme au pluriel) sera celle qui est plus facilement disponible et accessible et donc celle utilisée en premier, ce qui aboutit à une erreur de remplacement. Ce scénario développemental d'abord exploré par Morin (2003) puis par Chevrot, Dugua et Fayol (2005) est ici confirmé. La théorie qui peut prédire les erreurs des enfants avec autant de précision est ainsi le MBU.

Pallaud et Savelli (2001) ont analysé les erreurs langagières d'enfants d'âges variés en syntaxe, sémantique, morphologie et phonétique, et en ont comparé certaines avec des erreurs similaires chez les adultes. En ce qui concerne liaisons, chez les jeunes enfants, quelques erreurs de segmentation sont typiques de l'oral enfantin et sont absentes de la langue adulte. Les enfants semblent confondre *il y a* et *il a* et voulant faire une troncation en /ja/, ils font à la place une troncation en /la/. Ils en font de même avec « il y avait » qu'ils prononcent /lavɛ/ au lieu de /javɛ/. Les enfants semblent avoir remarqué qu'il y a troncation dans ce contexte, mais ils font la troncation avec un autre phonème, probablement parce qu'il est difficile pour eux de différencier *il y a* de *il a*. Excepté quelques cas uniquement présents chez les enfants, la plupart des erreurs se retrouvent au moins dans le type d'erreur (peut-être pas pour un item précis, mais pour ce genre de structure) chez les adultes. Celles-ci consistent principalement à faire de fausses liaisons, soit à rajouter un phonème de liaison pour éviter l'hiatus là où il n'en faut pas, soit à faire une liaison interdite, soit à utiliser le mauvais phonème, c'est-à-dire faire des cuirs et des velours. Aucune étude n'a fait état de liaisons obligatoires non effectuées chez les adultes ou chez les enfants de plus de 6-7 ans

Nous avons classé les erreurs enfantines de l'étude de Pallaud et Savelli (2001 :124-127) par âge de production (tableau 2.10.). On remarque qu'avec les liaisons obligatoires, il s'agit de cas particuliers : avec un h aspiré (*des-z-haricots, trois-z-hamsters, deux-z-héros*), avec une semi-consonne (*mon-z-oiseau*), et avec les enchaînements : *leur-n-anniversaire...*

4-5 ans	il était pas-t-encore né ;
6 ans	Souris qui-z-habitent ; des petits qui-z-embêtent les autres ; ceux qui-z-ont pas cantine ; je suis-t-allée me coucher ; on est-l-allé au cinéma ;
5-8ans 8 ans	ce qui est bien-z-en haut ; vous êtes-t-allées au Jouet-Club ;
9 ans	on lisait tout le mot-t-entier ; des-z-haricots ;
6-10-12 ans 8 et 12 ans	leur-n-anniversaire ; ton-n-hamster ;
11 ans	cinq-z-années ; et plein de petits-t-exercices ; trop-z-ancien ; deux-z-héros ;
âge non précisé	grande-t-épée ; mon-z-oiseau ; trois-z-hamsters ;

**Tableau 2.10. Erreurs de liaisons des enfants de 4-12 ans, d'après Pallaud et Savelli (2001 : 124-127)**

Dans tous les cas, les séquences comprenant des erreurs sont relativement peu fréquentes. On peut deviner que ce qui a « causé » les erreurs, ce sont les séquences fréquentes se trouvant en haut de la hiérarchie (tableau 2.11.) qui ont amorcé le début ou la fin d'une séquence, alors que celle-ci a été interrompue par un « intrus » qui ne se trouve pas à cette place dans le lexique usuel de l'enfant.

<b>Erreur</b>	<b>séquence la plus fréquente correspondante</b>
Souris qui-z-habitent	ils habitent / elles habitent
des petits qui-z-embêtent les autres	ils embêtent
ceux qui-z-ont pas cantine	ils ont pas
il était pas-t-encore né	il est encore / il était encore
je suis-t-allée me coucher	il est allé
on est-l-allé au cinéma	il allait / il est allé
vous êtes-t-allées au Jouet-Club	/vuzɛtale/ ils sont allés
on lisait tout le mot-t-entier	ce mot tout en entier, il est entier
cinq-z-années	des années / deux années / trois années / dix années
et plein de petits-t-exercices	un petit exercice
trop-z-ancien	très ancien
ton-n-hamster	ton ami, ton anorak
leur-n-anniversaire	mon anniversaire / ton anniversaire / ton anniversaire
grande-t-épée	grand effort / grand + nom sing
mon-z-oiseau	les oiseaux / des oiseaux
des-z-haricots	des + nom pluriel commençant par voyelle ou h muet
trois-z-hamsters	trois + nom pluriel commençant par voyelle ou h muet
deux-z-héros	deux + nom pluriel commençant par voyelle ou h muet

**Tableau 2.11. Erreurs enfantines, séquences en haut de la hiérarchie, Pallaud et Savelli (2001 : 124-127)**



L'acquisition des liaisons chez les enfants reflète un processus commun à tous les éléments phonétiques et toutes les composantes et structures linguistiques. Les études en acquisition d'une L1 montrent que la fréquence joue un rôle dans l'acquisition d'items, de structures et de représentations abstraites dans tous les domaines notamment : lexical, morphosyntaxique et phonologique. Plusieurs facteurs entrent en jeu dans l'acquisition des liaisons chez les enfants : l'usage du langage en réception permet à l'enfant de valider les séquences les plus fréquentes dans son lexique. Celui-ci, à son tour, utilise les séquences les plus fréquentes, ce qui lui permet de les ancrer et de construire des représentations abstraites.

L'enfant crée un réseau d'associations sur des bases lexicales ayant des implications morphosyntaxiques, mais aussi phonologiques et sociolinguistiques. Il est difficile de séparer les facteurs qui sont tous interdépendants : fréquence des phonèmes, fréquence des contextes possibles, fréquence des structures, fréquences des liaisons réalisées, fréquence du mot1, fréquence du mot2, fréquence de cooccurrence des mots, mais aussi saillance, opacité et simplicité d'une structure, ainsi que d'autres facteurs.

Les enfants plus âgés (de 11-12 ans) ont fait des erreurs avec des déterminants pluriels + h aspiré (*ton hamster, trois hamsters, des haricots*). Si l'on fait une recherche sur *Lexique* et que l'on fait une recherche dans *Film2(tout)*, on voit que les taux de cooccurrence des séquences en question sont les suivants : *ton hamster* : 6 ; *trois hamsters* : 0 ; *des haricots* : 92 ; *Est allé* : + de 1000 ; *Suis allé* : + de 1000.

La séquence *suis allé* a un taux de cooccurrence élevé et pourtant la séquence est produite avec une erreur par un enfant de 6 ans. On ne peut pas seulement considérer la fréquence de cooccurrence, il faut également considérer la productivité d'une structure, qui dépend elle-même du nombre d'items qu'elle englobe par rapport à ses concurrentes. Pour les formes verbales, les 3<sup>es</sup> personnes et les terminaisons en /t/ sont les plus fréquentes, ce qui pousse les locuteurs (enfants et adultes) à faire des liaisons avec /t/ (je suis-t-allée) au lieu

d'utiliser une consonne de liaison moins fréquente. Le MBU semble donc permettre d'expliquer les erreurs des enfants comme celles des adultes, mais aussi de prédire le développement langagier de l'enfant et les productions de liaisons des adultes.

Les études sur l'acquisition des liaisons des enfants L1 dont nous venons de parler font ressortir les tendances observées dans les autres études sur le MBU (Ellis, 2002 ; Tomasello, 2000a, 2000b). Dès qu'ils sont bébés, les enfants commencent à enregistrer les informations qu'ils reçoivent dans leur input, et utilisent un traitement statistique extrêmement efficace et fiable (Saffran, Aslin, et Newport, 1996). Après avoir appris les phonèmes les plus faciles d'un point de vue articulatoire, les enfants acquièrent les phonèmes les plus fréquents dans la langue ambiante (Monnin, Loevenbruck et Beckman, 2007) en même temps que les particularités des mots qui en déterminent leur utilisation, et ce de manière extrêmement sophistiquée (Ellis, 2002 ; Tomasello, 2000a, 2000b).

Dès deux-trois ans, les enfants apprennent les liaisons les plus fréquentes (les liaisons obligatoires) dans leur input (Chevrot et Fayol, 2000 ; Chevrot, 2002 ; Dugua (2002) ; Chabanal, 2003 ; Chevrot, Dugua et Fayol, 2005). Les différences dans l'input (qui dépend largement de celui des parents) se reflètent dans l'acquisition et les productions des enfants de tous les âges (Hoff, 2002 ; Nardy, 2003 ; Chevrot, Dugua et Fayol, 2005). Si les enfants sont privés d'input, ou que les informations ne peuvent pas être traitées « normalement » (problèmes d'audition d'accès aux informations, ou autres), leur cerveau ne peut pas enregistrer d'items, ils ne peuvent pas créer de classes et encore moins créer des catégories abstraites (Barlow et Kemmer, 2000). Ils ont ainsi des difficultés langagières allant de légers handicaps à une incapacité totale à maîtriser le langage.

Le processus d'acquisition de la liaison consiste à enregistrer des items, des séquences de mots, comme si les groupes de mots ne constituaient qu'un seul mot. Le nom, l'article et la liaison qui va avec l'article, ne forment qu'un. Cela permet d'apprendre le genre du mot, en

même temps que la bonne prononciation des séquences de mots et des liaisons correspondantes. D'où des erreurs consistant à rajouter l'article et la liaison qui l'accompagnent, même quand il ne le faudrait pas (*le navion*) (Dugua, 2006). Les enfants apprennent ensuite les liaisons facultatives les plus fréquentes (qui sont quand même moins fréquentes que les liaisons obligatoires) et enfin les liaisons facultatives moins fréquentes.

Il en est de même pour l'apprentissage de la lecture. Les enfants apprennent les mots les plus fréquents et les structures les plus simples, ils se servent ensuite de l'analogie pour reconnaître les mots écrits (Sprenger-Charolles, et al., 2003) et les exceptions. Pour la lecture de séquences avec des liaisons, le rôle de la fréquence est d'autant plus important qu'il en existe plusieurs types. Il faut les avoir entendues de nombreuses fois pour en avoir une mémoire auditive (fréquence d'écoute), et savoir comment elles se prononcent. Il faut les avoir vues pour les reconnaître visuellement, avoir des connaissances orthoépiques qui permettent de faire le lien entre graphie et phonie et enfin avoir pratiqué la lecture jusqu'à ce qu'elle devienne un automatisme (fréquence de pratique de la lecture).

Les enfants utilisent les structures les plus fréquentes, les plus saillantes et les moins complexes qui se trouvent en haut de la hiérarchie, qui sont le plus facilement et le plus rapidement accessibles. On retrouve ce phénomène dans tous les domaines : morphologie, syntaxe, lecture, écriture, etc., ce qu'on remarque dans les erreurs enfantines étudiées par Pallaud et Savelli (2001). En fait, le MBU permet non seulement d'expliquer le processus d'acquisition et les erreurs des enfants, il permet aussi de prédire les étapes d'acquisition et les erreurs apparaissant dans le processus d'acquisition (Dugua, 2005 ; Dufeu, et al., 2006).

Enfin, comme nous l'avons vu au chapitre 1, on observe chez les adultes L1 les mêmes phénomènes de nivellement et de simplification par analogie, le même système de traitement des informations selon les statistiques d'une banque de données qui dépend de l'input, et surtout de la fréquence et de la richesse de l'input (Langacker, 1987).

## 2. Les productions de liaisons des apprenants du français L2

Les études sur la production de liaisons d'étudiants du français L2 sont peu nombreuses. À ma connaissance, les seules qui portent sur la production d'apprenants anglophones sont celles de Mastromonaco (2000), Howard (2004) et Thomas (2002). Dans les trois cas, les productions de liaisons obligatoires sont très élevées et semblent se rapprocher de celles des francophones. Considérons plus en détail les résultats de chacune.

Mastromonaco (2000) a analysé les productions de liaisons obligatoires, facultatives et interdites d'anglophones de niveau intermédiaire. L'auteure inclut dans la catégorie des liaisons obligatoires certaines liaisons que nous considérons (ainsi que Thomas, 2002) comme facultatives très fréquentes : *c'est* + préposition, *c'est* + déterminant, adjectif ou adverbe, et adverbe + adjectif.

Mastromonaco (2000) a étudié les productions de deux groupes : un de 19 étudiants, et un autre de 30 étudiants (description pp.127-128). Chaque groupe avait un professeur différent, avait des exercices différents et les groupes ont reçu un enseignement à un an d'intervalle. Il était pertinent de les comparer pour voir s'il y avait des similarités en dépit des conditions légèrement différentes. Les étudiants étaient en 2<sup>e</sup> année d'université et suivaient un cours d'un semestre de français oral. Le pré-requis était soit d'avoir suivi le cours universitaire de français de 1<sup>e</sup> année, soit d'avoir suivi toute sa scolarité en école d'immersion. Les étudiants natifs ou ayant une maîtrise quasi-native du français n'avaient pas le droit de le suivre. Ils parlaient tous l'anglais et certains d'entre eux parlaient une autre langue en plus de l'anglais (ils étaient bilingues), mais il n'y avait pas d'autre précision sur la langue maternelle des sujets.

Les résultats (Mastromonaco 2000 :137 et 139) indiquent d'abord qu'il n'y a pas de différence significative entre les deux groupes dans les productions de liaisons obligatoires. La moyenne de tous les groupes pour toutes les tâches (lecture, parole spontanée et

description) était de 94%. Le groupe 1 en a produit 92.87% et le groupe 2 94.89%.

L'auteure résume les réalisations de liaisons obligatoires d'après les contextes syntaxiques (tableau 2.12.). Les liaisons sont réalisées à 100% dans 9 catégories, entre 95 et 97% dans 4 catégories et entre 90 et 93% dans 4 catégories et il y a aussi des pourcentages de réalisation très élevés pour les catégories le plus souvent considérées comme facultatives (c'est + ...). Ceci suggère une maîtrise parfaite du système de liaisons, à un niveau égal voire supérieur à celui des francophones, puisque dans l'étude de Malécot (1975), par exemple, les liaisons obligatoires sont réalisées à 96.9% par les francophones de France.

Contexte syntaxique	% de liaisons obligatoires réalisées
déterminant + adjectif	100%
déterminant + pronom	100%
clitique + clitique	100%
clitique + clitique	100%
préposition + adverbe	100%
préposition + pronom	100%
préposition + infinitif	100%
verbe + clitique	100%
expression figée	100%
préposition + nom	97%
clitique + verbe	96%
déterminant + nom	95%
préposition + déterminant	95%
adjectif + nom	93%
c'est + préposition	92%
clitique + verbe	92%
c'est + déterminant	90%
adverbe + adjectif	82%
adverbe + clitique	67%
c'est + adjectif	60%
c'est + adverbe	0%

**Tableau 2.12. Pourcentage de liaisons obligatoires d'après les contextes syntaxiques pour tous les locuteurs et toutes les tâches. Mastromonaco (2000 : 146)**

Mastromonaco conclut des productions des apprenants qu'ils maîtrisent bien les liaisons obligatoires et il est vrai que c'est ce que ces résultats indiquent :

The results show that there was a total of 955/1016 obligatory liaisons produced (94%). Only 61 liaisons were omitted. These findings were completely unforeseen, as it was expected that there would be a lower frequency of obligatory liaisons. The results reveal first of all, that obligatory liaison contexts occur frequently [...]. Secondly, the number of occurrences of obligatory liaison is high, since there is only a 6% error rate. This suggests that the notion of obligatory liaison seems to have been learned by FSL speakers. With the exception of the 61 omitted liaisons [...], obligatory liaisons do not appear to cause difficulties for language learners. Thirdly, these results reveal that the perception that L2 speakers have difficulties with liaison production is not entirely true. (Mastromonaco, 2000 : 137)

En ce qui concerne les liaisons facultatives, dans le premier texte (p. 300), on trouve: « veulent \_augmenter, des consommateurs\_en, mis\_en, naturels\_en. ». Dans le deuxième texte : « (p. 320) : « chercheurs\_américains, américains\_actuellement, actuellement\_employés, universités\_ou.

Mastromonaco (2000 : 168) indique que les locuteurs du groupe 1 produisent 17.87% de liaisons facultatives et les locuteurs du groupe 2 en produisent 10.80%. Il y a ici des différences significatives ( $t=5.286$ ,  $p(47df=.0000)$ ), soit une production moyenne de 13.59%. On peut attribuer ces différences de production entre les deux groupes au fait que le deuxième texte ne comprend qu'une seule véritable situation liante (*chercheurs américains*), les autres étant extrêmement rares, alors que le premier texte en comprend deux (*veulent augmenter, mis en*).

Mastromonaco (2000 : 168) en conclut que les apprenants maîtrisent les liaisons facultatives à un niveau similaire à celui des natifs en les comparant aux 21.7% de liaisons facultatives produites par les francophones de l'étude de De Jong (1989): « The frequency of optional liaisons in the speech of native speakers is just slightly higher than in the speech of FSL speakers. This suggests that since optional liaison is relatively similar in both groups of speakers, FSL speakers have also acquired the concept of optionality. »

En outre, il existe des différences de production entre les tâches (Mastromonaco, 2000: 169). Les locuteurs des deux groupes produisent 9.83% de liaisons facultatives dans la tâche de lecture, 13.37% dans la tâche de parole spontanée et 16.53% dans la tâche de description.

Le nombre restreint de liaisons facultatives (4 dans chaque texte) et le fait que les liaisons en question soient plutôt rares explique qu'il y ait un plus grand nombre de liaisons réalisées en parole spontanée et en description qu'en lecture. Les étudiants peuvent ainsi utiliser des structures plus fréquentes avec *être, pas, auxiliaire + participe passé*, etc.

On peut se demander pourquoi l'auteure en vient à la conclusion que les productions

de liaisons facultatives des francophones « sont légèrement plus élevées que celles des apprenants » et que les apprenants « ont aussi acquis le concept de ‘l’optionnalité’ ». En effet, en comparant les résultats de son étude et de celle de De Jong (1989), d’une part la différence paraît assez importante (13.59% comparé à 21.7%) et d’autre part, il est assez difficile de juger les connaissances des apprenants en ce qui concerne les liaisons facultatives, étant donné que plus de la moitié de ces contextes syntaxiques et lexicaux sont plutôt rares.

De plus, si l’on considère seulement les contextes susceptibles d’être produits : « veulent\_ augmenter, mis\_ en relief » dans le texte 1, et « chercheurs américains » dans le texte 2, on voit que ces contextes de liaisons sont assez fréquents et que les francophones les réaliseraient probablement, c’est-à-dire que leurs productions seraient sans doute nettement supérieures à celles des apprenants. Mais comme aucun francophone n’a lu ce texte, il est difficile de comparer les productions des apprenants du français L2 et les productions de De Jong (1989), puisque les contextes sont complètement différents. Autrement dit, il n’y a rien qui, selon nous, permette d’en venir à la conclusion que les apprenants maîtrisent les liaisons facultatives à un niveau similaire de celui des francophones.

Quant aux liaisons interdites, Mastromonaco (2000 : 188) considère qu’elles sont exceptionnellement produites, et qu’elles ne posent pas de problèmes pour les apprenants : « There was no difference among the two groups of speakers, since both produced a total of less than 4% forbidden liaisons. Forbidden liaisons did not cause major difficulties for either group. »

Plus précisément, il y a eu 0.79% de liaisons interdites dans la tâche de lecture (1 sur 126 contextes possibles), 7.14% dans la parole spontanée (4 sur 56 contextes possibles) et 3.86% dans la description (8 sur 207 contextes possibles), soit un total de 3.34% (13 sur 389 contextes possibles) de liaisons interdites réalisées (Mastromonaco, 2000 : 189).

Toutefois, ces résultats peuvent être interprétés autrement. Si les résultats sont plus

élevés en parole spontanée qu'en lecture (comme c'était le cas avec les liaisons facultatives), cela suggère d'abord que le texte contient peu de contextes possibles et que ces contextes sont très rares.

En effet, les contextes de liaisons interdites sont dans le texte 1: « question intéressante, et\_elle, emballages\_où, pourtant\_à », et dans le texte 2 : « et outre, américains\_actuellement, les universités\_où ». Il y a d'ailleurs peut-être eu des erreurs de codage car « américains actuellement, les universités où » étaient codés « 3 » (liaisons interdites) à la page 320 alors qu'à partir de la page 321, les mêmes contextes sont codés « 2 » (liaisons facultatives). S'il n'y a pas eu d'erreurs de codage et si l'auteur considère ces deux contextes comme étant des liaisons facultatives, alors cela signifie qu'il n'y a dans ce texte qu'un seul contexte de liaison interdite avec *et*. De tels contextes ont très peu de chances d'être réalisés d'une part parce qu'ils se trouvent entre des frontières syntaxiques (et ont une faible cohésion), d'autre part il y a eu des liaisons juste avant et juste après, et surtout ce ne sont pas des contextes fréquents. Bref, il est plutôt prévisible que les apprenants ne fassent pas de liaisons dans ces contextes, en lecture.

Quant à la parole spontanée et à la description d'image, si les étudiants maîtrisaient très bien les liaisons, ils devraient surgénéraliser les règles et produire des liaisons interdites, tout comme le font les francophones. On a vu dans les études de Boula de Mareüil et Adda-Decker (2002) et de Boula de Mareüil, Adda-Decker et Gendner (2003) que les francophones (de France) peuvent produire jusqu'à 10% de liaisons interdites selon les contextes. On a vu aussi que les enfants francophones continuent à produire des liaisons interdites et faire des erreurs de liaisons à 8 ans (Chabanal, 2007) et jusqu'à l'âge de 12 ans et peut-être même après (Pallaud et Savelli, 2001).

Le très haut taux de liaisons obligatoires et le faible taux de liaisons interdites suggèrent au premier abord que les apprenants ont acquis les liaisons à un niveau



francophone. Toutefois, 12 liaisons interdites réalisées (sur 470 contextes possibles) en parole spontanée et description est un chiffre très bas qui suggère plutôt un faible taux de surgénéralisation, et pose des questions quant aux représentations abstraites des apprenants.

Enfin, Mastromonaco (2000 : 213) estime que la prononciation de consonnes finales et le non-enchaînement des consonnes n'est pas problématique, puisque dans la majorité des cas les liaisons sont enchaînées : « The results from all speakers suggest that liaison consonant resyllabification did not prove to be problematic for these learners as most liaisons produced were resyllabified (93%). This confirms the conclusion previously mentioned that FSL speakers in both groups have acquired the notion of liaison. »

Or, le fait que 7% des liaisons ne soient pas enchaînées (Mastromonaco, 2000 : 275) suggère justement que les apprenants ne maîtrisent pas le système syllabique du français, et notamment les liaisons, à un niveau comparable aux francophones, qui ne produisent des liaisons non-enchaînées que dans des cas d'hésitations et de reprises, comme l'a remarqué Mallet (2008), les cas de liaisons volontairement non-enchaînées que mentionne Encrevé (1988) étant réservés à une minorité de personnages médiatiques (politiciens et journalistes).

Ainsi, les résultats de Mastromonaco (2000) sont à interpréter avec une certaine circonspection étant donné le contenu et les catégorisations des textes à lire, les méthodes d'analyse et l'interprétation des résultats.

Un autre chercheur qui a abordé les productions des apprenants du français L2 est Howard (2004, 2005, 2006). Il s'est surtout penché sur l'acquisition de la variation sociolinguistique, dont les liaisons. Dans son étude menée en Irlande (Howard, 2004), il n'y avait pas de texte à lire. Les 6 sujets ont étudié le français 5-6 ans à l'école, puis 2 ans à l'université. Au moment de l'étude, ils étaient en fin de deuxième année d'université et le français était leur spécialisation. Les anglophones étudiaient le français 7 heures par semaine, heures divisées entre un cours de langue comprenant une heure par semaine de conversation

et un cours de littérature ou de culture francophone. Howard (2004 :152) classe les étudiants de « *advanced instructed learners* » selon la définition de Bartning (1997).

L'étude portait sur les productions d'étudiants anglophones de liaisons obligatoires, facultatives et interdites d'après les contextes syntaxiques. Chaque entretien a duré une heure et se composait de parole spontanée d'après des questions variées de l'instructeur. Le nombre de questions portant sur différents sujets permettait de discuter de sujets variés.

Les résultats de l'étude étaient que les liaisons obligatoires ont été réalisées à 82.5% et les liaisons facultatives ont été réalisées à 8.2%. Howard (2004 :155) précise que ces productions sont éloignées des francophones qui font, selon Ashby (2003), 28% de liaisons facultatives. Vu le faible taux de liaisons facultatives, Howard en déduit que

[...] the learners first perceive obligatory contexts as contexts of use of the liaison in a way that variable contexts have yet to be acquired. [...].

In the case of our classroom learners, our findings have principally noted the relative underuse of obligatory liaisons, as well as the significantly lower usage of variable liaisons compared to the native speaker. [...] although they have been learning French for a number of years, and can be classed as 'advanced' instructed learners [...], the learners have significant progress to make on the acquisition of sociolinguistic variation. (Howard, 2004:158)

Dans une autre étude, Howard (2005) a étudié les productions de 12 étudiants en deuxième année d'université, ayant auparavant étudié le français 5-6 ans à l'école (tableau 2.13.). Au moment de l'étude, ils étaient en fin de deuxième année d'université et le français était leur spécialisation. Cette étude compare les productions d'un groupe de 6 étudiants étant resté au pays (1) et d'un autre groupe de 6 étudiants ayant passé une année scolaire en France (2). Le groupe 2 ayant passé une année en France a étudié avec les Français, mais n'a suivi aucun cours de langue cette année-là. Le groupe resté en Irlande a continué à étudier le français à l'université. Les liaisons obligatoires ont été réalisées à 82.2% par le groupe 1 et à 95% par le groupe 2 (ayant passé une année en France). Les liaisons facultatives ont été réalisées à 8.2% par le groupe 1 et à 12.7 % par le groupe 2. Les différences sont significatives. (Howard, 2005 : 11).

<b>Groupe 1</b>						
Apprenant	1	2	3	4	5	6
Article + substantif	96%	70%	70%	94%	95%	97%
Pronom sujet + verbe	100%	93%	100%	92%	100%	80%
Préposition monosyllabique	100%	82%	73%	78%	96%	75%
Adverbe monosyllabique	22%	25%	0%	100%	100%	33%
Expression figée	50%	40%	44%	50%	61%	50%
Adjectif + substantif	50%	0%	0%	-	0%	0%
Pronom objet	50%	0%	0	0%	0%	0%
<b>Groupe 2</b>						
Apprenant	1	2	3	4	5	6
Article + substantif	93%	100%	96%	100%	95%	97%
Pronom sujet + verbe	100%	100%	99%	97%	97%	100%
Préposition monosyllabique	100%	100%	94%	92%	100%	100%
Adverbe monosyllabique	20%	66%	0%	50%	60%	33%
Expression figée	100%	100%	100%	100%	85%	100%
Adjectif + substantif	-	33%	20%	29%	13%	0%
Pronom objet	-	33%	20%	29%	13%	0%

**Tableau 2.13. % Liaisons obligatoires selon le contexte syntaxique. Howard (2005 : 12)**

Il ressort de cette étude plusieurs tendances. Les étudiants ayant passé une année en France ont produit plus de liaisons obligatoires et facultatives dans toutes les catégories que ceux restés en Irlande. Les taux de liaisons réalisées entre un adjectif et un substantif, après un adverbe monosyllabique et avec un pronom objet restent faibles pour les deux groupes.

Les liaisons entre un article et un substantif sont produites dans une grande proportion, mais pas de façon systématique. Seuls 2 sujets (étant allés en France) sur 12 ont fait 100% de liaisons dans ce contexte.

En fait, on voit de la variation entre les sujets et entre les catégories syntaxiques. Seules les expressions figées ont été réalisées à 100% par 5 sujets sur 6 qui sont allés en France. On voit aussi que les catégories syntaxiques qui semblent être les mieux maîtrisées sont les structures les plus fréquentes (article + substantif, avec pronom, etc.) et celles qui ont été produites en plus faibles proportions (qui ont été acquises en dernier, ou qui sont même encore en cours d'acquisition) sont celles qui se trouvent dans des contextes moins fréquents, tels qu'après un adverbe monosyllabique ou entre un adjectif et un substantif, par exemple.

En ce qui concerne les liaisons facultatives, les deux groupes en produisent peu, et pour certaines catégories qui sont pourtant très fréquentes (souvent considérées comme

obligatoires) comme après *quand* (produite à 85% d'après le corpus des Français de Tours d'Ashby, 2003), seul un sujet a fait 17% de liaisons et tous les autres en ont fait 0%. Howard en conclut que leurs productions bien qu'assez élevées ne sont tout de même pas comparables à celles des francophones :

[...] bien que nos apprenants ayant fait un long séjour en France témoignent d'un taux de réalisation de la liaison plus élevé que les autres intervenants dans l'étude, ils ne se rapprochent tout d'abord pas du locuteur natif du point de vue de leur taux d'emploi général de la liaison. Dans un deuxième temps, pourtant, ils s'éloignent nettement du locuteur natif en raison de l'absence totale de liaison dans certains contextes syntaxiques. Ceci a pour conséquence d'engendrer un sentiment de non respect de toutes les contraintes sur l'emploi de la liaison chez le locuteur natif. (Howard, 2005 :15)

Enfin, après Mastromonaco (2000) et Howard (2005), l'autre chercheur qui a étudié les productions de liaisons des anglophones est Thomas (1998, 2002, 2004). Thomas (1998 : 543) indique que les liaisons sont problématiques à enseigner comme à apprendre pour les raisons suivantes. Tout d'abord, les liaisons sont extrêmement complexes dans leurs catégorisations comme dans leur utilisation. De plus, les francophones apprennent les règles complexes qui régissent les liaisons en les entendant. Ceci est difficilement possible pour les apprenants car, comme l'explique Thomas (1998), ils n'ont pas eu assez de contact avec des francophones et les liaisons sont rarement enseignées :

[...] [Les apprenants] sont tout simplement peu au courant des phénomènes de liaison, tout comme beaucoup d'autres apprenants du français langue seconde (ainsi d'ailleurs que leurs professeurs), la liaison étant rarement enseignée à cause de sa complexité.

[...] [La liaison] est souvent négligée dans les ouvrages de phonétique, classiques (Léon, 1965) ou modernes (Martin, 1996) et à peine abordée dans la plupart des manuels d'enseignement du français langue seconde. (Thomas, 1998 : 544)

En se référant à Huot (1978), Thomas (1998) précise que les élèves d'immersion ayant étudié le français pendant six années, bien qu'ayant habité en région francophone (près de Montréal), ne maîtrisent pas les liaisons aussi bien qu'on s'y attendrait, puisqu'ils ne font pas les liaisons obligatoires les plus fréquentes de façon systématique :

Pour l'ensemble des élèves, environ 20% des liaisons attendues n'ont pas été réalisées. Si c'est là tout ce qu'on peut obtenir dans une situation d'apprentissage quasi-idéale [...], on peut se demander ce qu'il en est des milieux moins favorisés. Le problème est aggravé par l'orientation pédagogique actuelle qui privilégie la communication au détriment du travail de structure. Les élèves finissent souvent par greffer sur des structures anglaises des mots français isolés et dénués de la cohésion nécessaire à la réalisation des liaisons. (Thomas, 1998 : 548)

Thomas (2002, 2004) a étudié les productions de liaisons de 87 sujets en 3<sup>e</sup> année de français. Il a comparé les productions de 39 sujets restés en Ontario à celles de 48 étudiants ayant passé une année scolaire en France. L'étude comprend en parole spontanée, une description de scène de rue, une description de l'université de l'étudiant et une autobiographie de l'étudiant au passé. Elle comprend aussi la lecture de deux textes : la lecture d'un texte sur la vigne et la lecture d'un passage d'*Astérix le Gaulois*. Les 87 sujets ont lu deux fois les deux textes (441 mots) ce qui fournit plus de 76 000 mots (Thomas, 2002 : 103). L'auteur compare ainsi les résultats obtenus par catégorie syntaxique (déterminant + nom, pronom personnel, etc.) aux études existantes, telles celle de Malécot (1975).

Les résultats obtenus montrent des taux de réalisation de liaisons très élevés : les étudiants anglophones produisent 91.1% de liaisons obligatoires (Ibid : 105) par rapport à 96.9% des francophones de Malécot (1975). Toutefois, Thomas (2002 : 105) précise que, même si ces productions paraissent proches de celles des francophones, il existe toutefois de grandes différences entre les productions de francophones et d'anglophones :

Bien que les résultats paraissent relativement proches, on peut voir en considérant les liaisons obligatoires mal ou non réalisées (8.9% contre 3.1%) que nos étudiants font en fait presque trois fois plus d'erreurs que les francophones dans ce domaine. Si on compare la liaison obligatoire aux autres difficultés analysées dans Thomas (1998), on constate qu'elle contribue à elle seule près de 20% de toutes les erreurs phonétiques. C'est là une proportion inquiétante pour un phénomène si répandu. (Thomas, 2002 : 105)

Thomas (2002) compare la production de certaines liaisons facultatives à celles des francophones de Ågren (1973). Les productions des anglophones sont beaucoup plus basses que celles des francophones. Par exemple, les francophones font la liaison avec *est* dans 97% des cas, alors que les anglophones ne la font que dans 66.2% des cas. Les anglophones ont aussi produit 9 liaisons interdites.

L'auteur (p. 108) en conclut que les étudiants maîtrisent peu les liaisons, qu'il faut prendre des mesures et enseigner la liaison de façon plus systématique et appropriée à leurs

besoins: « Une intervention pédagogique s'impose au niveau de la correction phonétique qui consiste à travailler d'abord la syllabation ouverte du français, qui s'applique à la fois à l'enchaînement et à la liaison, et à traiter ensuite cette dernière comme un cas particulier de resyllabation à droite, soumis à des contraintes entraînant une très grande variation. »

Thomas (2002 :117) considère ensuite les progrès réalisés par les deux groupes. En ce qui concerne les liaisons obligatoires, les étudiants ayant fait un séjour en France ont produit 86.7% de liaisons dans le pré-test (en début d'année scolaire) et 90.8% dans le post-test (en fin d'année scolaire). Les étudiants étant restés en Ontario ont produit 91.1% de liaisons dans le pré-test et 93.9% dans le post-test. En ce qui concerne les liaisons facultatives avec *est*, les étudiants étant allés en France ont produit 67.9% de liaisons dans le pré-test et 60.3% dans le post-test. Les étudiants étant restés en Ontario ont produit 65% de liaisons dans le pré-test et 73.8% dans le post-test.

De ces résultats, il apparaît d'abord que les deux groupes ont fait des progrès dans les productions de liaisons obligatoires ; le groupe expérimental (étant allé en France) a fait plus de progrès (+ 4.1%) que le groupe témoin (resté en Ontario) (+ 2.8%). Par contre, le groupe étant allé en France a régressé (- 7.6%) dans sa production de liaisons avec *est*, tandis que le groupe resté en Ontario a progressé (+ 8.8%).

Les étudiants ayant passé une année en France font (après leur séjour) moins de liaisons obligatoires (90.8%) et facultatives (60.3%) que ceux qui sont restés en Ontario qui en font 93.9% et 73.8%, parce que le groupe étant allé en France était plus faible dans sa production de liaisons au départ. Thomas commente cela dans les termes suivants :

Il est évident que huit mois d'études françaises ne peuvent pas changer grand-chose à la prononciation des étudiants avancés, déjà solidement établie avant la 3<sup>e</sup> année, et ceci même en milieu francophone. Ceci dit, les différences notées pour le groupe témoin vont toujours dans le sens, d'un progrès par rapport à la norme, ce qui suggère que l'enseignement dispensé cette année-là en contexte traditionnel ontarien a tout de même contribué indirectement à l'amélioration de la prononciation des étudiants. C'est rassurant... (Thomas, 2002 : 117)

Autrement dit, les étudiants font relativement peu de progrès parce qu'une bonne partie de la prononciation des apprenants est déjà acquise (et/ou fossilisée), et ce, même pour ceux qui ont passé une année en France.

Le tableau 2.14. résume ces études :

	francophones	Mastro-monaco	Howard Restés au pays (Irl.)	Howard Allés en France	Thomas Restés au pays (Can.)	Thomas Restés au pays(Can.)	Thomas Allés en France	Thomas Allés en France
		Moyenne 2groupes			Pré-test	Post-test	Pré-test	Post-test
étudiants au moment de l'étude		2 <sup>e</sup> année université Canada	2 <sup>e</sup> année université Irlande	2 <sup>e</sup> année université Irlande	3 <sup>e</sup> année université Canada	3 <sup>e</sup> année université Canada	3 <sup>e</sup> année université Canada	3 <sup>e</sup> année université Canada
Liaisons obligatoires	Malécot (1975) 96.9%	<b>93.32%</b>	82.5%	<b>95%</b>	91.1	<b>93.9%</b>	86.7%	<b>90.8%</b>
Liaisons facultatives	De Jong (1989) 21.7% Ashby (2003) 28%	13.9%	8.2%	12.7%				

**Tableau 2.14. Productions de liaisons d'anglophones d'après Mastromonaco (2000), Thomas (2002) et Howard (2005).**

Ces trois études indiquent des tendances similaires : les étudiants semblent maîtriser les liaisons obligatoires à un niveau proche des francophones. Les anglophones de Mastromonaco (2000) réalisent les liaisons obligatoires à 93.32%, les liaisons facultatives à 13.9% et les liaisons interdites à 3.34%. Les anglophones de Howard (2005) ont produit entre 82.5% et 95% de liaisons obligatoires, et ceux de Thomas (2002) en ont réalisé entre 86.7% et 93.9%. Autrement dit, les anglophones produisent, en moyenne, entre 90.8% et 95% de liaisons obligatoires, ce qui est très élevé et se rapproche des résultats des francophones (96.9% dans l'étude de Malécot, 1975).

Les étudiants de ces trois études avaient approximativement le même niveau (intermédiaire B2 d'après le CECR), étaient en 2<sup>e</sup> et 3<sup>e</sup> année d'université, étaient anglophones et avaient suivi des cours similaires (langue et littérature) avec l'approche communicative dans les cours de langue au lycée et à l'université.

Or, si ces chiffres paraissent d'abord très encourageants puisque les pourcentages de productions sont élevés, une étude plus approfondie fait ressortir plusieurs faits qui suggèrent des tendances différentes.

## 2.1. Analyse de ces études et interprétation des données

Ces études ne permettent pas, à notre avis, d'identifier les difficultés auxquelles font face les apprenants du français L2.

Les taux de réalisation très élevés des trois études suggèrent que les étudiants sont de niveau très avancé et parlent à un niveau proche de celui des francophones. Or, plusieurs phénomènes suggèrent le contraire. Mastromonaco (2000 :137) remarque à juste titre que : « These findings were completely unforeseen, as it was expected that there would be a lower frequency of obligatory liaisons. »

En effet, toute personne ayant enseigné le français aux anglophones de ce niveau (2<sup>e</sup>-3<sup>e</sup> année d'université) sait bien qu'ils ont des problèmes avec la syllabation, que ce soit la prononciation (et surtout la non prononciation) des consonnes finales, en particulier avec les voyelles nasales, les enchaînements et les enchaînements des liaisons. Ce qui est surprenant, c'est justement qu'ils réalisent un taux si élevé de liaisons.

D'ailleurs, Mastromonaco (2000:193) remarque certains problèmes en expliquant que les liaisons interdites pourraient être produites pour des raisons phonétiques : « [...] however, these liaisons may not have been produced for syntactic reasons, or as the result of failing to apply the rules of liaison. In fact, there may be a phonological explanation for their production, such as the nasal liaison consonant /n/ in *c'est une question intéressante*. »

Elle précise que les nasales sont particulièrement problématiques pour les apprenants :

A liaison in this sentence may be due to the nasal consonant and not to a liaison error [...]. Rather than produce a nasal vowel in some cases, the nasal consonant will be realized, and the preceding vowel undergoes denasalization. This phenomenon is found predominantly with the indefinite article *un*, which results in [un] or [œn], whether before vowel-or consonant-initial words (as in *un enfant* [unpətitãfã], speaker 26, spontaneous). Generally, it may be assumed that forbidden liaisons are not produced. Those that are produced are the result of factors such as the difficulty of nasal vowels. (Mastromonaco, 2000 : 195)

Mastromonaco (2000 :196) explique aussi que certaines liaisons interdites ont été faites parce que les sujets ont prononcé des consonnes finales qui devraient être muettes : « Of



the six cases of produced liaison, there were two types: liaison after *long* and before *haut* [...].»

Enfin, Mastro Monaco (2000 : 218) constate que les apprenants ont des difficultés à resyllabifier et à enchaîner les consonnes : « [...] Yet non-resyllabification was found in the speech samples of both groups for approximately 7% of the realized liaisons. Non-resyllabification was found to occur in all three tasks with very little difference among the tasks, indicating that this process may not depend on the nature of the discourse. [...] »

Thomas (2002 :106) remarque lui aussi des difficultés de (re)syllabification, de prononciation de consonne finale et des difficultés avec les voyelles nasales : « Notons que parmi les liaisons correctement réalisées, 8.5% l'ont été sans enchaînement (28.6% pour un \_ homme) ».

Thomas (2002 : 106) précise que ces non-enchaînements ne sont pas faits pour les mêmes raisons que les hommes politiques français étudiés par Encrevé (1988) : « [...] la liaison sans enchaînement reflète généralement chez nos étudiants une hésitation devant un mot difficile (ses\_ennemis) ou plus simplement une influence de la syllabification fermée caractéristique de leur langue maternelle, l'anglais. » et que les étudiants font des erreurs en particulier quand la consonne de liaison « n'est pas la même à l'écrit et à l'oral (d →/t) » (p.106).

De plus, il mentionne (p. 106) que certaines catégories sont réalisées en bien plus faibles proportions (adj. + nom) que les francophones. Les résultats de Howard (2005 : 8) vont aussi dans ce sens. Les taux de productions de liaisons obligatoires dans les contextes : adjectif + nom, après pronom objet et après un adverbe monosyllabique sont très bas y compris pour le groupe étant allé en France.

Autrement dit, les connaissances des étudiants en phonétique du français semblent limitées. Ils prononcent des consonnes finales muettes, prononcent des consonnes de liaison

sans les enchaîner et ils ont des difficultés avec les voyelles nasales, y compris avec le mot le plus fréquent : l'article *un* qui est mal prononcé (ou non-enchaîné) dans presque le tiers des cas ; ils ont des difficultés avec les consonnes de liaisons qui se prononcent différemment de l'écrit ; ils hésitent devant les mots et subissent l'influence de la syllabation fermée de l'anglais ; enfin, ils produisent beaucoup moins de liaisons dans plusieurs catégories pourtant très fréquentes...

Malgré tout cela, ils produisent quand même entre 91 et 95% de liaisons obligatoires !? On ne peut guère penser que seule une minorité plus faible rencontre les difficultés phonétiques ci-dessus mentionnées, produit une faible proportion de liaisons, et fait baisser la moyenne (qui serait autrement supérieure à 95%) de la majorité au niveau plus élevé dont la plupart maîtrise très bien les liaisons. La raison en est que les étudiants moins avancés produisent en parole spontanée et en description d'image un discours plus court avec moins de contextes liants, et les seuls contextes qu'ils produisent (les plus fréquents) sont le plus souvent réalisés avec des liaisons, ce qu'on voit dans les annexes des auteurs. Ne produisant pas ou peu de contextes liants, leur production ne fait pas baisser la moyenne.

Comment se peut-il donc qu'avec toutes ces difficultés et ces connaissances limitées, les productions globales de liaisons soient si élevées dans les trois études ?

### **2.1.1. Les tests utilisés**

Tout d'abord, les tests utilisés ne permettent pas d'évaluer les connaissances et compétences des apprenants en ce qui concerne différentes catégories syntaxiques et phonétiques, ni d'évaluer les contextes un peu plus rares. Les textes que les sujets anglophones devaient lire étaient des textes authentiques qui n'ont pas été conçus pour tester la réalisation des liaisons dans la lecture à haute voix. Les textes en question étaient : une publicité pour l'étude de Mastro Monaco (2000), un passage d'un texte sur la vigne et un

passage d'Astérix dans l'étude de Thomas (2002). (Dans l'étude de Howard, il n'y avait pas de texte à lire).

Si ces textes ont le mérite d'être naturels et authentiques, ils sont parfois courts (les deux textes de l'étude de Mastro Monaco font 129 et 155 mots), et surtout non « concentrés » en liaisons, puisqu'il est rare dans des textes « naturels » de trouver plusieurs liaisons dans une même phrase ou plusieurs liaisons à la suite. Vu qu'on peut difficilement demander aux sujets de lire plusieurs pages à haute voix, cela restreint les études à un nombre limité de contextes. Les textes n'abordent que certains points de certaines catégories syntaxiques et phonétiques, les contextes les plus fréquents et les plus « faciles » de liaisons obligatoires et sont donc les plus susceptibles d'être maîtrisés.

Dans les trois études, il y avait des tests de parole spontanée ou semi-guidée (description d'image). Dans de tels exercices, les étudiants risquent de n'utiliser que les liaisons les plus fréquentes. Les liaisons plus rares susceptibles de ne pas être réalisées ne sont pas produites du tout, et dans les textes à lire elles sont également absentes.

Pour mieux comprendre les inconvénients d'utiliser la parole spontanée pour contrôler les connaissances des apprenants, on peut comparer les productions de parole spontanée des apprenants du français L2 à celle des francophones. On voit alors que les jeunes enfants francophones utilisent des groupes de mots ou tronçons préfabriqués, qu'ils ont le plus entendus (Tomasello, 2000b : 157), de manière répétitive ;

Les adultes francophones utilisent, eux-aussi, de nombreuses formules préfabriquées (Erman et Warren, 2000), ce qu'on remarque notamment chez les commentateurs sportifs ou dans les ventes aux enchères, lorsque les commentateurs doivent parler très vite. En parole spontanée, les locuteurs ont tendance à répéter plusieurs fois les mêmes séquences, séquences d'ailleurs les plus fréquentes (Boula de Mareuil et Adda-Decker, 2002).

Si l'on observe les productions en parole spontanée des sujets de l'étude de Mastromonaco (2000 : 299-349), on voit que les apprenants ont un vocabulaire plus limité et font plus de répétitions que des locuteurs natifs et utilisent de façon répétitive les formules préfabriquées les plus simples (faute de meilleures connaissances en vocabulaire).

On observe une forte tendance à répéter les mêmes mots et expressions dans les productions orales transcrites dans l'étude de Mastromonaco (2000 : 299-349). Parmi ces nombreuses répétitions, on note par exemple que le sujet 1 répète dix fois « il y a » et cinq fois « il a »; ces tronçons de la plus grande simplicité sont répétés un grand nombre de fois ce qui suggère que le sujet a des connaissances fort limitées en français.

De même, le sujet 2 répète cinq fois « on voit », cinq fois « il y a », trois fois « on remarque » et deux fois « on regarde » ; le sujet 44 répète deux fois « il y a », deux fois « je pense », quatre fois « ils sont », quatre fois « peut-être » et deux fois « ils ont ». Le sujet 46 répète cinq fois « il y a » et trois fois « et puis » ; le sujet 47 répète cinq fois « ils sont » et le sujet 49 répète aussi quatre fois « ils sont », etc.

Puisqu'ils répètent souvent les mêmes mots et groupes de mots, cela signifie que le fait de ne compter que le nombre de liaisons ou enchaînements réalisés est un facteur dont il faut tenir compte dans l'interprétation des résultats. Par exemple, le sujet 1 qui a répété dix fois *il y a* et cinq fois *il a*, aura probablement fait les enchaînements dans ces deux cas, soit un total de quinze (bien que ceux-ci ne soient pas vraiment étudiés ici). On ne peut pourtant pas en déduire que le sujet maîtrise les enchaînements parce qu'il en a fait quinze, mais seulement qu'il maîtrise les enchaînements dans ces deux contextes très fréquents.

Ainsi, la parole spontanée peut être trompeuse pour une étude de ce genre, car les sujets utilisent les structures les plus fréquentes, qui sont sans doute les structures qu'ils connaissent le mieux, les prononçant probablement correctement (avec la liaison adéquate). Néanmoins, on ne peut pas généraliser à tout un système les quelques exemples produits qui

ne sont que les cas les plus fréquents. Si un apprenant dont les connaissances du français sont limitées réalise 80 ou 90% de liaisons obligatoires en parole spontanée, cela ne veut pas dire que ses productions se rapprochent des francophones, mais cela indique seulement que les quelques liaisons présentes dans sa parole (parmi des contextes très limités) ont été réalisées, et ce, précisément parce que sa parole est limitée et répétitive.

En effet, les apprenants L2 évitent les structures moins fréquentes qu'ils maîtrisent moins bien, tout comme les locuteurs natifs les évitent aussi. Mais puisque les apprenants L2 en maîtrisent beaucoup moins que les locuteurs L1, cela veut dire qu'ils vont aussi en éviter beaucoup plus. Comme l'indiquent par exemple Lightbown et Spada (2006 : 82) : « [...] learners sometimes avoid using certain features of language that they perceive to be difficult for them. This avoidance may lead to the absence of certain errors, leaving the analyst without information about the learner's developing interlanguage. [...] »

C'est la raison pour laquelle la parole spontanée en L2 ne peut pas entièrement tester les compétences générales en liaisons, mais seulement la performance dans les séquences les plus fréquentes.

D'une part, le vocabulaire et les structures sont limités, d'autre part certaines structures se répètent et d'autres sont absentes. Il est ainsi prévisible que les seules structures qui apparaissent en parole spontanée chez des locuteurs L2 se trouveront dans les contextes les plus fréquents : le contexte [art. + nom] et avec les noms les plus fréquents; d'où l'utilité d'un test de lecture qui permette de déterminer plusieurs contextes.

Puisque dans l'étude de Mastro Monaco, par exemple, la parole spontanée ne permet pas de rendre compte des structures moins fréquentes, et que le texte à lire ne teste que les liaisons obligatoires les plus fréquentes telles que les liaisons dans les contextes les plus simples et fréquents tels que [art. + nom] (*les Américains, les Américains*<sup>33</sup>, *des ambitions*,

---

<sup>33</sup>La séquence les Américains est répétée deux fois dans le texte

*les universités, les entreprises*) et entre certains adjectifs fréquents + nom (*son intention*), contextes les plus susceptibles d'être maîtrisés, ni la parole spontanée, ni la lecture ne vont pouvoir évaluer avec précision les contextes que maîtrisent (ou maîtrisent moins bien) les apprenants. C'est ce qui explique les taux de production si élevés.

De plus, dans le texte à lire de l'étude de Mastromonaco (2000), certains contextes sont exactement les mêmes, tels que « les Américains » répétés deux fois, ou presque identiques, tels que pour la préposition *en* dans deux contextes (*en octobre, en Europe*).

Précisons d'ailleurs que parmi les liaisons avec les déterminants, d'après l'étude de Malécot (1975 : 164), c'est avec l'article pluriel que la liaison est faite en plus grand nombre (99.6%) par les francophones, par opposition aux adjectifs qui ne sont réalisés qu'à 93.2%.

Dans la même étude de Malécot (et à la même page), parmi les prépositions, la préposition *en* est celle qui est la plus sujette à la liaison (ainsi que la préposition *chez*). Elle est réalisée à 100%, alors qu'elle n'est réalisée qu'à 98.8% avec *dans* et à 85.7% avec *sans*. Autrement dit, les items testés, sont eux-mêmes les plus fréquents de leur catégorie. L'étude teste donc les items les plus fréquents à l'intérieur des structures les plus fréquentes.

Ainsi, à notre avis, les résultats des études de Mastromonaco (2000), Thomas (2004) et Howard (2004) reflètent seulement la maîtrise des liaisons avec les items les plus fréquents dans les catégories les plus fréquentes, ce qui est déjà beaucoup pour des non-natifs. D'ailleurs, même à l'intérieur de ces catégories, et parmi ces items les plus fréquents, leurs compétences sont déjà discutables puisque les liaisons ne sont réalisées qu'entre 91% (moyenne des deux groupes) dans l'étude de Mastromonaco et parmi ces liaisons, 7% n'ont pas été enchaînées, 93% pour Thomas et 95% pour le groupe de Howard qui a passé une année en France. En d'autres termes, si même les catégories les plus fréquentes, avec les items les plus fréquents (par exemple avec la préposition *en*) présentent déjà quelques

difficultés, cela suggère que le système n'est pas complètement maîtrisé et laisse présager des lacunes avec des structures plus complexes et moins fréquentes.

Pour déterminer avec plus de précision si l'acquisition des liaisons dans certains contextes phonétiques et syntaxiques est similaire à celles des francophones, il importe de tester plus que les items les plus fréquents et les structures les plus fréquentes et de considérer plusieurs classes ou sous-classes (adjectifs singuliers, pluriels, avec nasales, etc.), et de comparer exactement la même chose : le même contexte phonétique, syntaxique et lexical, et le nombre de contextes liants produits.

Si on ne compte que les pourcentages réalisés (et non le nombre de contextes possibles), lorsque des locuteurs produisent des structures très simples, ne comprenant par exemple que le contexte [dét. + nom], cela indique justement que leur niveau de français n'est pas très élevé, puisqu'ils ne produisent aucun autre contexte liant. Le nombre de contextes liants des apprenants L2 devrait être comparé au nombre de contextes liants chez les locuteurs L1, comme le fait Hieke (1984).

Seulement si le nombre et surtout la variété de contextes est comparable, peut-on comparer les pourcentages de productions d'apprenants L2 et de locuteurs natifs. Il faudrait alors comparer à la fois les contextes syntaxiques, phonétiques et fréquentiels et en dresser une liste très précise, ce qui serait difficile et long à réaliser en production spontanée. Comme cela prend trop de temps, les études comparent seulement des catégories syntaxiques générales de type [dét. + nom].

Mastromonaco a précisé qu'aucun des sujets de 2<sup>e</sup> année n'avait un niveau natif, ce que l'on constate en effet dans les productions des apprenants en parole spontanée et en descriptions (que l'on trouve dans les annexes, pp. 299-349), indiquant un niveau maximum B2 et même parfois B1. Les phrases des locuteurs sont simples et courtes, elles comprennent de nombreuses fautes de grammaire et de vocabulaire et de très nombreuses répétitions. C'est-

à-dire que même si le niveau B1-B2 est un niveau déjà élevé pour des apprenants, c'est un niveau quand même éloigné de celui des francophones.

Une analyse des productions orales permet d'expliquer en partie le taux élevé de liaisons obligatoires. Tout d'abord, pour la parole spontanée et les descriptions, la simplicité des phrases et des structures limitent fortement les contextes de liaisons. Par exemple, le sujet 3 (pp. 301-302), dans sa très courte production, n'a produit aucun contexte de liaison, le sujet 2 (p. 300) n'en produit qu'un seul : *mon avis*, et le sujet 5 en description (p. 304) n'en produit que deux : *des yeux, un oiseau* dont la liaison n'a pas été faite avec *un oiseau*.

Autrement dit, moins les étudiants sont de niveau avancé, plus leurs productions sont courtes et simples, moins ils produisent de contextes de liaisons et plus les rares contextes liants sont susceptibles de produire des liaisons puisque ce sont les structures les plus simples et les plus fréquentes. Par contre, plus les locuteurs ont un niveau élevé, plus leurs descriptions sont longues, précises et complexes, plus ils produisent de contextes liants. Si ces locuteurs sont de niveau plus élevé, ils sont plus à même de produire des liaisons. Ainsi, les étudiants les moins avancés ne produisent aucun contexte ou très peu, ce qui ne fait pas baisser les résultats, (puisque'on ne considère pas le nombre limité ou l'absence de contextes liants). Comme il n'y avait pas ou peu de contextes liants pour commencer, le peu qu'ils produisent est correctement réalisé pour la plupart. Quant aux locuteurs les plus avancés, ils réalisent la plupart des liaisons dans des contextes liants plus variés. Finalement, ce sont les locuteurs les plus avancés qui risquent de faire baisser légèrement les moyennes en produisant des structures complexes avec des items moins fréquents.

De plus, les erreurs des locuteurs empêchent elles aussi les sites possibles de liaisons et reflètent par ailleurs le manque de connaissances de la langue française de nombreux sujets ; par exemple le sujet 5 (p.304) dit *au haut* (au lieu *d'en haut*), empêchant ainsi une liaison (interdite) possible ; un locuteur a dit *à côté du oiseau*, trois locuteurs ont dit *une*



*oiseau*, et un locuteur a dit *une petite oiseau* (p.159). Dans ces cas, il n'y a pas de contextes liants. On ne considère donc pas qu'il y a eu des erreurs de liaisons, ou des non-réalisations de liaisons, puisqu'il n'y avait pas de contexte liant à proprement dit. Il s'agit pourtant d'erreurs, et c'est justement à cause de ces erreurs que les liaisons n'ont pas pu être produites. Cette méthodologie est donc problématique, car elle permet d'obtenir des chiffres très élevés qui indiquent ce que les locuteurs maîtrisent, mais n'indiquent pas ce qu'ils ne maîtrisent pas.

### **2.1.2. Classement des liaisons, catégorisations et définitions**

En plus du type de test utilisé, les classements et catégorisations peuvent aussi changer les résultats. L'interprétation des résultats dépend de plusieurs facteurs notamment des définitions et des critères de départ de l'étude. Selon qu'on considère des liaisons comme obligatoires ou facultatives, on n'obtiendra pas les mêmes résultats. Les définitions et les catégorisations des liaisons n'étant pas clairement définies, certains chercheurs considèrent les mêmes liaisons comme obligatoires, et d'autres chercheurs considèrent les mêmes liaisons comme facultatives.

Les résultats très élevés des trois études semblent indiquer une maîtrise des liaisons à un très haut niveau, proche des natifs. Rappelons les données sur les productions de locuteurs natifs que nous avons vues précédemment. Selon Malécot (1975), aucune liaison n'est réalisée à 100% sauf avec la préposition *en*. Toujours selon ce dernier, la liaison est réalisée à 98% avant un nom (entre déterminant et nom), à 99.5% après tous les articles, à 98.6% après un pronom personnel sujet, à 98.8% après une préposition, et à 93.2% après un adjectif. Pour Boula de Mareüil, Adda-Decker et Gendner (2003), les liaisons sont réalisées à 95.3% après un déterminant, à 90.5% après une préposition monosyllabique et à 72.5% entre un adjectif et un nom.

De plus, moins on considère de sous-classes dans la catégorie « liaisons obligatoires », plus les classes considérées ont de chances d'être les plus fréquentes et donc « de gonfler » les résultats. Par exemple, si on considère les contextes [adj. qualificatif + nom], [aux. + verbe] comme appartenant à la catégorie des « liaisons facultatives », ou que ces catégories sont absentes, cela signifie qu'il ne reste plus dans la catégorie liaisons obligatoires que [art. + nom], [pron. +...], [...+ pron.] et [prép. monos. +...]. Dans ce cas, le pourcentage de liaisons obligatoires va être plus élevé que si un plus grand nombre de catégories y appartenait. Quant au pourcentage de liaisons facultatives, il va lui aussi être plus élevé, si on y inclut des liaisons obligatoires ou quasi-obligatoires.

Selon la définition qu'on donne à une liaison « bien formée », on n'obtiendra pas les mêmes résultats non plus. Si l'on considère que la prononciation d'une consonne latente suffit pour être considérée comme « liaison réalisée » (ce qui est le cas dans l'étude de Mastromonaco), on n'obtient pas les mêmes résultats que si l'on considère que pour qu'une liaison soit (correctement) réalisée, elle doit être prononcée avec la bonne consonne de liaison, (comme c'est le cas dans l'étude de Thomas, 2002), doit être enchaînée au mot<sup>2</sup> et que les phonèmes précédant et suivant la consonne de liaison doivent être correctement prononcés (comme dans notre étude).

Cela pourrait expliquer en partie les résultats du pré-test de Thomas (2002), dont les étudiants ont un niveau comparable à ceux de Mastromonaco (dans le pré-test les étudiants de Thomas sont en début de 3<sup>e</sup> année et dans l'étude de Mastromonaco ils sont en 2<sup>e</sup> année). Les étudiants obtiennent des résultats légèrement inférieurs (86.7% et 91.1%) à ceux de Mastromonaco (93.32%). Thomas restreint un peu plus que Mastromonaco les cas de liaisons en précisant que la consonne de liaison doit être correctement prononcée.

Les trois chercheurs considèrent que même si les consonnes ne sont pas enchaînées, elles rentrent tout de même dans la catégorie des liaisons produites. Nous ne partageons pas

cette interprétation de ce qui constitue une liaison correctement produite. Nous considérons que pour qu'une liaison soit correctement produite, elle doit être produite telle qu'un locuteur natif la produirait dans une même situation : elle doit être prononcée avec la bonne consonne de liaison et être enchaînée au mot<sup>2</sup>. Nous considérons que sinon, ce n'est pas une liaison, mais simplement la prononciation de la consonne finale du mot<sup>1</sup>. Autrement dit, notre définition est plus stricte que celle des études précédentes, et pour cette raison nous enlèverions d'ores et déjà les 7% de liaisons non-enchaînées aux 93.3% des liaisons de Mastromonaco et 8.5% aux 86.7% et 91.1% des liaisons de Thomas.

### **2.1.3. Les éléments testés**

Vu le type de tests utilisés, nous considérons que les études discutées n'analysent pas assez de structures syntaxiques et phonétiques, et que parmi les structures testées (article + nom, préposition + nom), non seulement les items sont parfois les mêmes (par exemple, *les Américains, les Américains, en octobre, en Europe* dans les textes de Mastromonaco), mais aussi que les items sont les plus fréquents et les plus « faciles » à réaliser.

Pour déterminer avec plus de précision si l'acquisition des liaisons dans certains contextes phonétiques et syntaxiques est bien similaire à celles des francophones, il importe de tester plus que les items les plus fréquents et de considérer plusieurs classes ou sous-classes telles que adjectifs singuliers, pluriels, avec nasales, expressions figées, etc. Les francophones utilisent une plus grande variété de structures, si bien que si on regroupe adjectifs possessifs, démonstratifs, interrogatifs, indéfinis, articles définis et indéfinis dans une même catégorie [dét. + nom], si l'on compare les productions de [dét. + nom] des locuteurs L1 et L2, en réalité on ne compare pas la même chose si les locuteurs L2 ont seulement utilisé la structure [art. + nom].

D'autre part, le problème de la syllabation, bien que brièvement abordé, n'a pas véritablement été étudié, alors qu'il est crucial dans la production de liaisons. Si on ne teste pas les enchaînements séparément, on ne peut pas déterminer si les problèmes de syllabation sont dus à des difficultés à resyllaber ou si c'est un problème particulier aux liaisons. Une étude parallèle sur les enchaînements, en plus des liaisons, permet de considérer avec plus de précision les questions de syllabation.

D'autres études se sont penchées sur la segmentation ou le découpage syllabique, mais pas vraiment sur les enchaînements. Cutler et al. (1986, 1989), Beaudoin (1998), Steele (2001) et Golato (2002) ont comparé la segmentation de francophones et d'anglophones bilingues. Ces études en segmentation testent plutôt comment un sujet découpe un mot tel que bateau : « bat-eau, bat-teau ou ba-teau » (Beaudoin, 1998 : 362), perçoit un mot ou un groupe de mots ou comment il peut apprendre à syllaber dans une deuxième langue, ou comment un locuteur bilingue aborde le découpage syllabique.

Saunders (2007) a étudié principalement la question de la chute du schwa et quelques cas isolés d'enchaînements dans la perception des apprenants. Les résultats ont montré que les étudiants avaient des difficultés à reconnaître les groupes de mots dans lesquels les schwas n'étaient pas prononcés et avaient également des problèmes à reconnaître des cas d'enchaînements et d'effacements.

Stridfeldt (2002) a analysé la perception des enchaînements d'étudiants suédois apprenant le français. Leurs difficultés indiquent que si les étudiants ne peuvent pas reconnaître les mots enchaînés, ils ne les enchaînent probablement pas eux-mêmes, car production et compréhension sont intimement liées. On voit ici que l'enchaînement est problématique pour de nombreux groupes linguistiques, tels que les locuteurs du suédois ou de l'anglais.

Les études de Mastro Monaco (2000), Thomas (2002) et Howard (2005) n'ont pas analysé la production d'enchaînements comme phénomène indépendant (ce n'était pas leur objectif) et encore moins comme complément à la liaison en français L2.

En outre, l'interférence de la langue maternelle est un facteur qui mérite une attention particulière. C'est pourquoi, si l'on veut déterminer les compétences des apprenants, on devrait aussi tester ce qui est difficile pour les locuteurs d'une L1 donnée, notamment (mais pas seulement) les enchaînements pour les anglophones. Il serait aussi utile d'évaluer les langues maternelles des apprenants, notamment pour les sujets bilingues.

## **Discussion sur les productions d'apprenants du français L2**

Nous pensons que les pourcentages élevés de productions de liaisons indiquent seulement les structures et items les plus fréquents que maîtrisent les locuteurs L2, à cause des classements et catégorisations, des méthodes et tests utilisés, mais aussi à cause des principes d'acquisition des liaisons selon le MBU, d'après les conditions d'apprentissage des étudiants comparées à celles des enfants francophones dont nous avons discuté au début de ce chapitre.

Puisque les étudiants n'ont pas reçu d'instruction formelle en phonétique du français, puisque l'approche pédagogique se concentre sur la communication plus que sur la précision, puisqu'ils ont eu un input trop restreint pour leur permettre d'internaliser les règles, et que l'input reçu n'est pas toujours constitué de séquences bien formées si les enseignants ne sont pas toujours de langue maternelle, nous présumons que les apprenants ne peuvent pas maîtriser les règles abstraites qui gouvernent les liaisons et les enchaînements.

Cela prend presque 10 ans aux enfants francophones qui se trouvent dans des conditions maximales, qui entendent et pratiquent le français entre 10 et 12 heures par jour, chaque jour avant d'arriver à un niveau comparable à celui des adultes. Il paraît impossible que des étudiants apprenant le français quelques heures par semaine, puissent arriver aux

mêmes résultats étant donné leurs conditions d'apprentissage. Puisque la qualité et la quantité de l'input que les apprenants ont reçu sont forcément inférieures à celles des francophones, les apprenants ne peuvent pas maîtriser les liaisons au même niveau que les francophones.

Alors que Purcell et Suter (1980) analysent les *Predictors of pronunciation accuracy*, on peut prédire les productions de liaisons, et surtout les NON productions de liaisons, ou les liaisons mal formées, d'après certains facteurs tels que la fréquence des items (mots1, mots2 et cooccurrence des mots1 et mots2), la fréquence et la productivité des structures, la fréquence des consonnes de liaison, le type de formation reçue, le nombre d'années d'étude du français, la langue maternelle, et d'autres facteurs sociaux propres aux locuteurs.

Si les étudiants n'ont pas été amenés à se pencher sur la forme, ici, sur la prononciation et plus particulièrement sur les liaisons, les enchaînements et la syllabation, il est peu probable qu'ils aient pu avoir appris à resyllaber par eux-mêmes et à produire les liaisons et enchaînements comme les natifs. Ils n'ont pas entendu les items et les structures un nombre suffisant de fois pour pouvoir construire des représentations abstraites des lois régissant ces aspects phonétiques. Nous discuterons plus en détail de tout cela dans notre étude expérimentale.

Krashen (1985) a souligné le rôle de l'input dans le processus d'acquisition. Comme l'indique Lauret (2007 : 35) en ce qui concerne l'approche communicative : « Les promoteurs de cette approche n'ont pas véritablement discuté du rôle de la prononciation dans l'enseignement des langues, ni développé de stratégies pour enseigner la prononciation de façon communicative. L'objectif n'est pas de tenter une prononciation la plus proche possible de celle des natifs, mais seulement que la prononciation n'entrave pas la capacité à communiquer ».

Autrement dit, puisque la non-réalisation de liaisons et enchaînements n'entrave presque jamais la communication avec les autres apprenants ou avec l'enseignant (sauf cas

d'opposition singulier / pluriel, de type : *il arrive / ils arrivent ; leur ami / leurs amis*), il est peu probable que ces points aient été abordés de façon formelle en classe, et qu'il y ait eu des corrections systématiques. En outre, les enseignants sont rarement formés à la précision et à la correction d'éléments phonétiques.

Si correction ou explication il y a eu, celles-ci n'ont jamais été enseignées de manière systématique puisque cela irait à l'encontre de l'approche communicative, qui consiste surtout à être capable de communiquer et non à étudier la langue de manière analytique (ce qui a été confirmé dans les réponses à notre questionnaire). Ce qui serait donc surprenant c'est que les étudiants arrivent à maîtriser des règles qu'ils n'ont jamais étudiées formellement portant sur des structures et des items qu'ils n'ont pas assez entendus pour les avoir appris naturellement.

Comme nous l'avons vu dans le chapitre 1, les enfants francophones apprennent d'abord les items isolés les plus fréquents, ou plus précisément les séquences les plus fréquentes se composant de deux ou trois mots (Dugua, 2005) de type *l'âne, l'avion, les amis* ; puis apprennent des séquences un peu moins fréquentes. Après avoir acquis un assez grand nombre d'items et de séquences (dont le nombre reste à déterminer), ils finissent par internaliser des règles. Ils sont alors capables de produire des énoncés qu'ils n'ont jamais entendus. Ce n'est qu'à ce stade que le niveau de maîtrise de certaines structures peut être considéré comme étant à un niveau natif adulte.

D'après les études mentionnées précédemment, les productions enfantines se rapprochent de celles des adultes à 6 ans pour les liaisons obligatoires et à environ 10-11 ans pour les liaisons facultatives et les erreurs de liaisons. Or, d'après les études en acquisition L1 et L2, d'après le MBU et d'après le niveau des étudiants de l'étude (niveau intermédiaire B1-B2 d'après le CECR), à ce stade de leur apprentissage, et vu leurs conditions d'apprentissage, les étudiants anglophones ne peuvent sans doute produire que quelques items individuels (lexicalisés) - ce qu'on remarque d'ailleurs dans leur production spontanée - structures et

items qu'ils ont le plus souvent entendus, notamment dans le contexte le plus fréquent [art. + nom].

Dans le cas d'élèves anglophones, il paraît évident que leur fréquence d'écoute et que le nombre de liaisons entendues est beaucoup plus restreint que pour les francophones majoritaires. Aucun des sujets des études n'a étudié le français dans une école francophone. Le français entendu dans les écoles n'est pas toujours parlé par des natifs. Le français entendu en classe vient des autres élèves et surtout des enseignants de français souvent anglophones.

Or, le modèle donné joue un rôle primordial, comme nous l'avons vu chez les enfants francophones (entre autres Hoff, 2002 et Dugua, 2005, 2006). Si les enseignants n'étaient pas francophones, ils ont probablement eux-mêmes produit moins de liaisons que des francophones majoritaires et ont sans doute fait des erreurs de syllabation, notamment de prononciation de consonnes finales muettes.

De plus, les contextes d'écoute du français que les élèves rencontrent en salle de classe sont plus réduits dans le nombre d'heures comme dans la variété des situations. D'après ces conditions d'apprentissage et l'input plus réduit que celui des francophones majoritaires, les étudiants du français L2 devraient réaliser les liaisons obligatoires, les liaisons facultatives et les enchaînements dans un nombre beaucoup plus restreint que les francophones majoritaires et aussi dans un nombre plus restreint que les francophones minoritaires.

Les recherches en acquisition dans le cadre du MBU, les recherches sur l'acquisition des liaisons des enfants francophones et les résultats obtenus dans les études des anglophones que nous avons mentionnées s'opposent. Il faut 6 ans aux enfants francophones avant de maîtriser les liaisons obligatoires et jusqu'à 10-11 ans pour maîtriser les liaisons facultatives et interdites, alors qu'ils ont entendu les séquences bien formées des milliers de fois. Cela prendrait 20 ou 30 fois moins longtemps à des non-natifs n'ayant pas toujours entendu des séquences bien formées (étant donné le nombre de séquences entendues et le nombre d'heures



d'écoute et de pratique de la langue) pour arriver à des résultats similaires ? Il y a ici une inadéquation.

Nous pensons que ce décalage entre les résultats escomptés d'après les conditions d'apprentissage et l'input reçu et les résultats obtenus dans les trois études discutées (une maîtrise quasi native des liaisons) est dû aux types de tests utilisés, aux textes à lire et à la définition de ce qui constitue une liaison « correctement réalisée » et à l'interprétation des résultats. Enfin, nous pensons qu'il est impératif de comparer ce qui est comparable : la lecture d'un même texte par des natifs et des non-natifs.

## **CHAPITRE III**

### **MÉTHODOLOGIE ET HYPOTHÈSES**

## 1. La problématique

Comme nous l'avons montré dans le chapitre 2, les études portant sur la production de liaisons des étudiants anglophones sont peu nombreuses et ne cernent pas complètement les difficultés auxquelles ils font face. Le taux de productions de liaisons est élevé parce que ces études analysent seulement les contextes les plus fréquents en production libre et en lecture.

D'autre part, les études sur les productions d'enchaînements des non-francophones en français sont, à ma connaissance, inexistantes. Mastromonaco (2000), Thomas (2002) et Howard (2004) ont indiqué que les étudiants avaient des problèmes à resyllaber (c'est-à-dire à réaliser les enchaînements), mais ce n'étaient que des remarques, en passant, puisque les études ne portaient pas à proprement parler sur ce sujet.

Comme mentionné au chapitre 2, Mastromonaco (2000 : 218) remarque que 7% des liaisons n'ont pas été enchaînées et Thomas (2002 : 106) que parmi les liaisons enchaînées, « 8.5% l'ont été sans enchaînements ». Ces pourcentages ne prennent en compte que les cas rencontrés dans les liaisons, et non les cas d'enchaînements (en dehors des liaisons) non réalisés, ainsi que les cas de consonnes muettes prononcées. Il serait donc utile de faire une analyse plus approfondie sur les enchaînements et les liaisons non-enchaînées.

Dans ces études, en parole spontanée et description d'images, les étudiants les plus faibles produisent peu de contextes liants. Comme la façon de calculer les résultats consiste à déterminer les contextes réalisés par rapport aux contextes possibles, ceux qui n'en ont produit aucun ne font pas baisser le taux de production de liaisons, tandis que les seules productions réalisées sont le plus souvent maîtrisées, et celles qui sont moins maîtrisées pourraient avoir des non-enchaînements. Le type de test pourrait ainsi refléter le niveau des étudiants les plus avancés et ceux de niveau intermédiaire, car ce sont eux qui ont produit le plus de contextes liants et ont eu les productions les plus longues.

D'après le MBU et d'après le niveau des étudiants des études mentionnées, notre hypothèse est qu'à ce stade de leur apprentissage, la plupart des apprenants ne peuvent produire qu'un certain nombre d'items individuels (lexicalisés) les plus fréquents, ceux qu'ils ont le plus souvent entendus. Nous pensons qu'ils maîtrisent partiellement les structures syntaxiques les plus fréquentes [art. + nom] et ce uniquement dans les séquences les plus fréquentes avec les mots<sup>1</sup> et les mots<sup>2</sup> les plus fréquents, et non de manière systématique, comme c'est le cas pour les francophones.

Les productions orales et la lecture des textes ne cernent pas les véritables enjeux des liaisons. Ce n'est pas parce que certains locuteurs (les plus avancés) produisent correctement certaines liaisons en parole spontanée, alors que certains n'en produisent pas du tout, et une bonne partie en produisent peu (1, 2 ou 3 dans l'étude de Mastromonaco), faute de contexte liant, que l'on peut en déduire, comme le fait Mastromonaco, que les étudiants maîtrisent le système complexe des liaisons. La seule façon d'évaluer leurs connaissances est de les amener à utiliser des structures qui leur sont moins familières, ce qui n'est possible que dans un exercice de lecture contenant des contextes plus variés. Mais pour que cet exercice soit utile, il faut que le texte cerne de nombreuses difficultés, catégories et sous-catégories.

## **2. Méthodologie**

Nous allons expliquer notre choix de test et de méthodologie, expliquer pourquoi certaines méthodologies ont été choisies dans les études précédentes et pourquoi nous pensons qu'elles peuvent être problématiques. Après avoir évalué les avantages et inconvénients de chacune, nous expliquerons pourquoi nous avons préféré le test de lecture pour évaluer la production de liaisons et d'enchaînements des apprenants anglophones.

Plusieurs méthodes existent pour évaluer la production orale des apprenants L2 : la

parole spontanée ou guidée, la parole semi-guidée (description d'images), la répétition et la lecture à haute-voix. Les études précédemment citées ont toutes utilisé la parole spontanée et la parole semi-guidée (description d'image, et/ou questions soigneusement préparées pour orienter les réponses des sujets) et deux d'entre elles ont utilisé la lecture en supplément à la parole spontanée et semi-guidée. Les résultats obtenus dans ces études sont similaires, et ce, quel que soit le type d'activité.

## **2.1. La parole spontanée**

Elle permet à l'apprenant L2 de s'exprimer librement et de fournir aux évaluateurs des informations sur les productions en situation authentique de communication. Les productions en parole spontanée permettent de contrôler certaines compétences acquises. Pour que des structures ou des règles soit bien utilisées, elles doivent avoir été entendues et produites un certain nombre de fois jusqu'à ce qu'elles deviennent immédiatement accessibles.

Si certaines structures sont bien utilisées dans la production spontanée, c'est le plus haut niveau de compétence. Si on peut utiliser correctement des règles phonétiques dans la parole spontanée, en principe, on peut également les utiliser correctement (sauf erreur de parcours) dans la parole semi-guidée, dans la répétition et dans la lecture.

Par contre, on peut être capable de répéter ou de lire un texte en utilisant correctement certaines règles phonétiques, sans pour autant être capable de les utiliser en parole spontanée. En effet, la parole spontanée requiert la mise en œuvre de nombreuses compétences simultanément : grammaire, vocabulaire, prononciation. La parole spontanée peut être considérée comme étant la parole la plus « naturelle » et la plus représentative de la réalité linguistique. D'où le commentaire de Green et Hintze (1990 : 62) : « We therefore take as axiomatic that any empirical study of sandhi in French must be based on authentic, connected discourse. »

Toutefois, si la production orale (tout comme la production écrite) permet à l'apprenant d'utiliser ce qu'il maîtrise bien, ce type d'activité lui permet aussi d'éviter toute structure ou vocabulaire qu'il maîtrise moins bien, et d'éviter à *fortiori* les structures inconnues, que ce soit des structures syntaxiques, lexicales ou phonétiques.

Lorsque les apprenants L2 ont été encouragés très tôt à s'exprimer dans la langue cible (ce qui est généralement le cas avec l'approche communicative), ils ont recours à des stratégies de compensation telles que l'utilisation de synonymes, de périphrases, de circonlocutions et d'évitements. Si c'est une stratégie efficace pour pouvoir s'exprimer assez librement, ceci peut aussi représenter un problème pour tester certains points linguistiques.

Comme l'explique Lefrançois (2005 : 429), les francophones eux-mêmes évitent ce qui peut créer des erreurs à l'écrit : « Producing a text is easier than correcting one, because avoiding potential errors is more convenient than solving imposed problems. » Si les francophones eux-mêmes évitent les structures difficiles, ceci est d'autant plus le cas pour des apprenants L2 qui maîtrisent évidemment moins bien la langue que les natifs.

De même, en ce qui concerne la production orale, les apprenants L2 ont tendance à utiliser des structures simplifiées et préfabriquées qu'ils ont fréquemment entendues et répétées. Ceci est naturel étant donné que leur vocabulaire est plus limité que celui des francophones.

Beeching (2001), Ovtcharov, Cobb et Halter (2006) (cités dans Thomas, 2008) ont comparé le vocabulaire des francophones et des anglophones. Il ressort de ces études utilisant le logiciel *Vocabprofil* (dont nous discuterons plus tard) que les natifs utilisent beaucoup plus de mots « rares » de la catégorie hors-liste (HL) que les apprenants L2. Autrement dit, les apprenants L2 utilisent plus de mots simples et moins de mots rares que les locuteurs natifs.

De plus, en production orale, les apprenants L2 ont tendance à répéter encore plus souvent qu'à l'écrit les mêmes mots, puisqu'ils n'ont pas le loisir de chercher des mots dans le

dictionnaire, de cliquer sur la fonction « synonyme » ou même de réfléchir aux formes utilisées.

En effet, nous avons déjà discuté des répétitions des étudiants dans l'étude de Mastromonaco (2000 : 299-349) au chapitre 2. L'exemple de Thomas (2008 : 596) extrait d'une production semi-guidée d'un étudiant de 3<sup>e</sup> année universitaire (de niveau intermédiaire /avancé de niveau B2) ayant passé une année scolaire en France est tout aussi révélateur.

Le mot *mec* est répété 5 fois et le mot *attaché* est répété 2 fois en 43 mots (les erreurs ont été conservées telles que citées) : « Un mec... ben, sa tête n'est pas attachée. Ce mec a un mec dans son poche et ses pieds sont pas attachés ... son jambe. Et un mec avec un queue. Un squelette qui vend des journaux, avec un mec en haut. »

Cet extrait est, à notre avis, représentatif des productions des étudiants (de 2<sup>e</sup>-3<sup>e</sup> année) des trois études dont nous venons de discuter, car la langue utilisée se rapproche des productions de Mastromonaco (en dehors de l'utilisation du français familier). Le passage contient 3 fautes de genre en 2 lignes, des structures simples et répétitives, indiquant une maîtrise limitée de certains aspects lexicaux et morphosyntaxiques du français, même après une année passée en France. Le passage en question (certes plus court que ceux des études), contient deux liaisons facultatives avec *pas* et une liaison interdite avec *en haut*, et donc aucune liaison obligatoire.

Étant donné que les étudiants ont appris certains mots et expressions qu'ils répètent de manière quasi-automatique, ces mots et expressions sont sans doute bien prononcés, car ce sont des mots fréquemment entendus et fréquemment produits (ici, le mot *mec*). Si les apprenants prononcent certains mots et expressions de façon similaire aux locuteurs natifs, cela peut donner l'impression que leur français est proche de celui des locuteurs natifs, mais cela ne veut pas dire qu'il le soit vraiment.

Par exemple, dans une étude portant sur les liaisons en parole spontanée, si la méthode

consiste à considérer les contextes possibles et les comparer aux contextes réalisés, on peut arriver à un pourcentage de réalisation des liaisons très élevé. S'il y a 10 contextes possibles et 9 contextes réalisés, cela veut dire qu'il y a 90% de liaisons (qu'il était possible de faire) réalisées. En analysant de telles productions, on pourrait avoir l'impression que l'apprenant maîtrise les liaisons parce qu'il peut en avoir produit un assez grand pourcentage par rapport aux contextes de liaisons possibles. On pourrait conclure qu'un apprenant qui, dans la parole spontanée, réalise 90% de liaisons (possibles) maîtrise bien les liaisons.

Les dangers d'une telle méthode et de telles interprétations sont que les chiffres ne prennent pas en compte si c'est toujours le même type de liaison qui est utilisé, avec les mêmes mots ou type de mots ou expressions. Par exemple, si le seul contexte dans le groupe nominal est le contexte [dét. + nom], et avec des mots très fréquents (*hommes, enfants*) et d'autant plus avec des articles pluriels (catégorie de déterminant la plus réalisée par tous les francophones car la plus fréquente, comme on le voit notamment dans l'étude de Malécot, 1975), les sujets devraient forcément faire des liaisons.

Ou bien, si dans le contexte [verbe + GN], les seuls cas liants sont des liaisons avec le verbe *être*, qui est une liaison facultative très fréquente, à la limite d'être obligatoire (Mastromonaco la considère d'ailleurs comme obligatoire), les sujets pourraient utiliser souvent les mêmes structures et items, et si c'est le cas, ils feront la liaison de nombreuses fois, car ils devraient connaître et maîtriser ces structures fréquentes qu'ils répètent souvent. C'est ce qu'on voit en effet dans la production du sujet 1 de Mastromonaco (2000 : 299) : « Je suis une [...] elle est en train de lire [...], elle est en train de parler [...], la femme est en train de lui demander, [...] est accroché [...]. »

Les mêmes mots ou expressions sont considérés de la même façon que s'ils étaient des mots différents, alors qu'en réalité on peut considérer qu'il s'agit d'une seule et même liaison, avec « être » et surtout « elle est en train », qui est presque lexicalisé. Ainsi, ces quelques



liaisons ne suffisent pas à représenter la classe [verbe + nom], [verbe + prép.], [verbe + part. passé]. Toutefois, puisque les productions sont classées par catégories syntaxiques, et que ce contexte appartient à cette catégorie, alors les nombreuses liaisons font rapidement monter les taux de production de liaisons. On retrouve aussi ce même phénomène avec les productions de francophones, mais dans leurs productions les contextes syntaxiques et phonétiques sont plus variés, ainsi que leur lexique, limitant les répétitions et étendant les contextes de liaisons possibles à des contextes moins fréquents dans lesquels la liaison est moins « obligatoire ».

Si l'on prend en compte uniquement les pourcentages réalisés des productions orales (spontanées et semi-guidées), on arrive à des taux de liaison élevés. Autrement dit, il est ainsi possible de déterminer surtout les contextes qui paraissent maîtrisés, qui sont d'ailleurs probablement plutôt lexicalisés (*est un, est une*, etc.), mais pas les catégories qui n'ont pas été produites. C'est ainsi que les informations sont incomplètes.

En effet, dans la parole spontanée des études de Mastromonaco (2000), Howard (2005) et Thomas (2002), de nombreuses sous-catégories, telles que adjectifs numéraux, adjectifs avec nasales, semi-consonnes, expressions figées, etc. n'ont pas été produites et l'on ne peut donc pas savoir si elles sont maîtrisées. De nombreux contextes ne sont présents ni dans la parole spontanée, ni dans la lecture. Il est ainsi difficile d'analyser des catégories quand celles-ci sont incomplètes. Il est certes impossible d'étudier tous les contextes dans une seule enquête; par contre il est possible d'étudier de nombreux contextes, ce que nous avons tenté de faire dans notre étude.

En plus de cela, se rajoute l'interprétation des données. Si une liaison n'apparaît qu'une seule fois dans une catégorie, on ne peut guère en tirer de conclusions. Par exemple, si l'on trouve 3 liaisons réalisées sur 3 possibles dont une entre un déterminant et un nom, une avec une préposition monosyllabique (100%), et une avec un adjectif antéposé fréquent (100%), on ne peut pas en déduire que l'apprenant L2 maîtrise parfaitement les liaisons

obligatoires, et ces sous-catégories, parce qu'il a réalisé, au total, trois liaisons obligatoires les plus fréquentes avec 100% dans chaque catégorie. Par contre, si un sujet réalise 45 liaisons obligatoires sur 50 (comme dans notre étude), alors c'est déjà un peu plus révélateur.

En fait, si ce sont les seuls (trois) cas trouvés en parole spontanée et qu'il n'y a pas d'autres cas de liaisons possibles (et celles-ci devraient pourtant être nombreuses), cela indique peut-être que la parole est simple, les phrases courtes et que c'est plutôt une indication du manque de maîtrise globale du français, plutôt qu'une indication de maîtrise, même si le locuteur maîtrise en effet les trois liaisons produites. Les cas qui demandent des structures un peu plus difficiles (pronoms COD, COI, certains pronoms indéfinis plus rares de type aucun, tout, etc.), expressions figées, mots commençant avec un H etc., seront probablement absents du discours.

Autrement dit, oui, les erreurs sont significatives, mais le manque d'erreurs l'est aussi. On ne peut pas se tromper avec ce qu'on ignore. Si l'on ne connaît aucune expression figée, aucun mot commençant par un h aspiré, qu'on ne sait pas utiliser les pronoms et qu'on préfère donc répéter des groupes nominaux avec [dét. + nom] plutôt que d'utiliser des pronoms, qu'on ne sait pas utiliser les adjectifs antéposés à la bonne place et qu'on hésite avec les chiffres, on ne va pas faire d'erreurs de liaisons dans ces catégories puisqu'on ne va pas les utiliser du tout.

Un autre problème avec les tests de production spontanée est d'isoler les facteurs. Il est parfois difficile de déterminer si ce sont les contextes phonétiques, syntaxiques, sémantiques, morphologiques, ou « fréquentiels » qui sont en jeu. Par exemple, si un apprenant prononce mal *un* et qu'il le prononce comme *une* /yn/ ou un mélange hybride de *un* et *une* : /œn/, ce que font très souvent les étudiants (Mastromonaco, 2000 : 219), on ne sait pas si c'est la prononciation qui est en cause, ou si c'est un problème de genre et donc une faute grammaticale. En outre, on ne peut pas considérer que c'est une liaison non plus (sous

prétexte que le /n/ a été prononcé), si les apprenants prononcent le /n/ de la même façon lorsque le mot suivant a une initiale consonantique. Cela indique seulement qu'ils prononcent une consonne finale dans tous les cas et non seulement en cas de liaison.

Par contre, on peut clairement déterminer dans un test de lecture que c'est une erreur de prononciation, puisque le mot est écrit et le locuteur ne doit pas choisir le genre. Il reste à déterminer si c'est un problème de prononciation des nasales, ou si c'est un problème de prononciation fautive de consonnes finales généralisable à toutes les consonnes, ce qui semble être le cas dans l'étude de Mastromonaco (2000 : 219-334) qui appelle les consonnes de liaison non enchaînées *Temporally anchored consonants*. Les étudiants ont tendance à prononcer de nombreuses consonnes finales muettes.

Enfin, la parole spontanée analyse plus la performance que les compétences de l'étudiant. Si la performance démontre une compétence certaine, au moins dans les séquences produites correctement, un manque de performance n'indique pas forcément un manque de compétence. Lorsque les apprenants L2 parlent de façon spontanée, ils se concentrent sur le message au détriment de la précision et ceci est particulièrement le cas pour la prononciation.

Comme le mentionne Lauret (2007) :

Il est difficile de se concentrer sur ce qu'on dit en même temps que sur la façon dont on le dit, en particulier dans une langue étrangère. On sait aujourd'hui que l'activité cérébrale est notablement différente quand il s'agit d'apprécier auditivement la forme d'un message ou d'accéder à son sens. L'entraînement doit donc porter sur la perception auditive et la répétition pour développer ces capacités avant de chercher à les transférer dans des énoncés à visée communicative. (Lauret, 2007 : 97)

Il est possible que les étudiants connaissent des règles de prononciation et soient capables de les utiliser s'ils se concentrent uniquement sur la prononciation, la lecture et/ou la répétition, mais qu'ils n'en soient pas capables s'ils se concentrent sur le message, en parole spontanée. Un étudiant de haut niveau peut hésiter parce qu'il ne se souvient pas d'un mot ou du genre de ce mot, et ne pas faire de liaison, alors qu'avec un autre mot, il l'aurait faite. Ce n'est pas parce qu'un apprenant n'a pas fait la liaison dans un cas particulier qu'il ne maîtrise

pas les liaisons de cette catégorie. En effet, Mastromonaco (2000 : 215-216) remarque que les étudiants ont tendance à ne pas resyllaber les mots quand ils se trouvent face à des mots plus longs et/ou inconnus.

De plus, il est possible qu'en parole spontanée, les apprenants n'aient pas produit certains contextes liants et certaines liaisons, simplement parce qu'ils n'ont pas pensé à utiliser certaines structures, alors qu'ils en sont pourtant capables. Ainsi, l'absence de certains contextes ne nous renseigne ni sur leur compétence, ni forcément sur leur manque de compétence.

La parole spontanée est un excellent outil pour tester les connaissances générales de vocabulaire, de grammaire, certains aspects phonétiques et autres compétences linguistiques, mais n'est, à notre avis, pas l'outil le plus fiable pour tester les compétences précises des apprenants en ce qui concerne les liaisons et les enchaînements dans différents contextes syntaxiques, phonétiques et fréquentiels, car il n'y aura pas assez de contextes assez variés et plus rares (h aspirés, etc.) pour déterminer avec précision les compétences de l'apprenant.

## **2.2. La parole semi-guidée**

La parole semi-guidée est proche de la parole spontanée, mais elle permet de mieux orienter les productions des étudiants, grâce aux questions et aux sujets de discussion (Howard 2004 : 153). Par exemple, dans l'étude de Thomas (2002), les sujets devaient faire leur autobiographie en utilisant le passé, ce qui favorisait l'étude du contexte [aux. + part. passé]. La parole semi-guidée permet ainsi d'obtenir des résultats plus précis que la parole spontanée (libre).

Si un sujet est capable d'utiliser correctement une structure dans la parole spontanée, c'est généralement qu'il maîtrise cette structure. Toutefois, si une structure est sur le point d'être maîtrisée, mais ne l'est pas tout à fait, le locuteur risque de l'éviter s'il ne se sent pas

assez à l'aise pour l'employer. Il arrive aussi que le locuteur utilise une structure dans un contexte particulier, si bien qu'on ne sait pas si c'était lexicalisé ou si le sujet maîtrise réellement l'élément testé.

C'est la raison pour laquelle la parole semi-guidée permet de tester une forme de parole quasi-spontanée, tout en permettant à l'évaluateur de diriger le sujet sur des points plus précis. L'instructeur peut éliciter des situations telles que la visualisation de films ou d'images, ou bien demander au sujet de continuer une phrase lorsque le début lui est donné.

Par exemple, dans les études de Mastromonaco (2000) et Thomas (2002), les sujets doivent décrire une image sur laquelle il y a certains indices, ce qui orientera le sujet dans ses commentaires. Plus la personne peut donner de détails, plus les descriptions sont longues, plus cela montre les connaissances, notamment au niveau lexical. On voit dans l'étude de Thomas (2008) que les descriptions des étudiants ayant passé un an en France sont plus longues que celles des étudiants restés au Canada, pour le même temps d'élocution (plus de mots par minute).

La parole semi-guidée permet de tester des éléments plus précis que la parole spontanée. Par exemple, dans l'étude de Thomas (2002), il y avait deux productions semi-guidées : pour l'une il fallait décrire une image, et pour l'autre il fallait parler de son université comme pour une publicité. On peut ainsi tester la maîtrise du contexte [dét. + nom], s'il y a plusieurs images avec sur une un enfant, un homme, sur une autre plusieurs personnes (*ils, elles*) et que les personnages des images ou photos font des actions avec des verbes commençant par des voyelles (*arriver, aimer...*). Cela réunit alors les avantages de la parole spontanée tout en pouvant mieux cibler certains contextes. Ce genre de test est efficace pour tester les connaissances lexicales et pour tester des contextes phonétiques plus restreints tels que [art. + nom], ce qui était le cas ici.

Ce genre de test présente pourtant quelques inconvénients. D'une part, il peut y avoir des erreurs de parcours, comme pour la parole spontanée; l'apprenant peut aussi utiliser d'autres groupes de mots, ou hésiter et faire autre chose que ce qui était prévu. Par exemple, si l'on veut tester la production de liaisons entre le pronom et le verbe, si le sujet se trompe de pronom et dit *il leur aime*, il est alors difficile de déterminer s'il maîtrise bien la liaison des pronoms s'il ne maîtrise pas les pronoms eux-mêmes. Dans la parole semi-guidée, comme dans la parole spontanée, de nombreux facteurs entrent en jeu, ce qui complique les tests et les interprétations.

De plus, il est assez difficile dans un tel exercice de tester un grand nombre de contextes, car cela prend beaucoup de temps pour éliciter les productions; cela peut prendre au sujet quelques secondes ou millisecondes entre chaque réponse pendant qu'il réfléchit. Les pauses, hésitations, reprises, temps utilisé par l'instructeur et pour la réflexion limitent le temps et donc les contextes, qui se réduisent souvent à une seule sous-catégorie. Ou alors, le test est très long comme dans l'étude de Howard, où les interviews duraient une heure par personne. Si le test est long, le sujet risque d'être fatigué, de se lasser des questions et du test et de ne pas fournir sa meilleure performance, surtout à la fin du test.

Les auteurs des études précédemment citées ont utilisé la parole guidée et semi-guidée car ils voulaient tester les contextes les plus fréquemment utilisés : [dét. + nom], [pron.+ verbe], [prép. + art.], et [aux.+ part. passé] qui seraient forcément présents dans ces productions. Mais si cette approche permet de mieux évaluer la parole naturelle, elle présente l'inconvénient majeur de restreindre l'étude à des contextes assez limités. Puisque nous souhaitons évaluer des contextes plus précis et plus variés de liaisons obligatoires, facultatives, interdites et d'enchaînements avec des restrictions de temps, la parole spontanée (ou semi-guidée) prendrait trop de temps et ne pourrait pas tester tous les points que nous voulons étudier.

### **2.3. La répétition**

La répétition est plutôt employée pour tester si les apprenants sont capables de bien répéter c'est-à-dire d'imiter, et s'ils sont capables de différencier certaines informations, certains phonèmes, intonations ou autres éléments segmentaux ou suprasegmentaux. L'un des critères des bons apprenants est leur capacité à bien imiter la langue cible (voir Suter, 1976 ; Suter et Purcell, 1980 ; Stapp, 1999 ; Yates, 2003).

Les tests de répétition sont une excellente façon de tester les capacités d'imitation. Alors que les autres méthodes testent surtout la compréhension, l'expérience, les connaissances et les compétences, cette méthode teste plutôt certaines aptitudes individuelles en reconnaissance comme en production.

On emploie souvent cette méthode pour tester les capacités des apprenants à entendre, discriminer et reproduire des éléments prosodiques, tels que l'intonation et le rythme (Neufeld, 1979 et 1980). Les tests de répétition peuvent parfois donner de meilleurs résultats (Neufeld, 1979 et 1980) que les tests de prononciation en parole spontanée pour déterminer certains aspects phonétiques tels que le degré de précision et de rapprochement de la parole native.

Toutefois, pour ce qui nous intéresse (les liaisons et enchaînements), il ne s'agit pas seulement de savoir si les apprenants sont capables d'imiter, mais plutôt s'ils sont capables d'internaliser des règles phonétiques et de produire des liaisons et enchaînements dans différents contextes. Les apprenants utiliseront la répétition lors des exercices de pratique en classe, mais pas pour le test.

## 2.4. La lecture d'un texte à haute voix

Champagne-Muzar et Bourdages (1993) proposent une typologie de l'enseignement de la prononciation, dont discute Lauret (2007 : 134-135) à partir des activités les plus « simples » consistant à sensibiliser les étudiants à certains aspects de la prononciation et les entraîner à discriminer et identifier des éléments phonétiques à des activités plus complexes de production dirigée et production spontanée. D'après cette typologie, et selon les auteurs, être capable de lire un texte à haute voix est une compétence qui se situe à mi-chemin dans l'acquisition.

Ce genre de test présente aussi plusieurs avantages : la lecture met en œuvre des compétences contrôlables et représentatives ; on peut déterminer les contextes à étudier et en avoir plusieurs dans chaque catégorie et l'on peut déterminer avec plus de fiabilité les contextes maîtrisés ou ceux qui créent des difficultés.

La lecture est, d'un côté, pour un apprenant L2, un exercice plus facile que la parole spontanée, dans la mesure où il n'a pas à se soucier de la grammaire et des éléments rentrant en compte dans la parole spontanée. De plus, un locuteur peut mieux se concentrer sur ses compétences phonétiques, puisque celles-ci ne sont pas entravées par les obstacles du contenu, du message, et des formes syntaxiques et sémantiques.

S'il y a une erreur, on sait alors que c'est une erreur phonétique et non grammaticale. Comme le précise Lauret (2007 : 159) : « [...] [La lecture est] une activité à préparer si l'on veut obtenir une image la plus proche possible de la compétence en prononciation / oralisation à partir de l'écrit. La préparation permet de porter l'attention sur la forme et non sur le fond, ce qui se traduit souvent par une nette amélioration de la performance. »

Il est possible de lire un texte en français, même si on n'en comprend pas tous les mots. Dans ce cas, on remarquerait d'ailleurs les véritables connaissances phonétiques (les représentations abstraites), plutôt que les connaissances lexicales ou grammaticales.



Par contre, la lecture présente aussi quelques inconvénients. Premièrement, elle active inconsciemment les mécanismes phonétiques de lecture de la langue maternelle, comme le soulignent par exemple Vendelin et Peperkamp (2005) et Detey et Nespoulous (2007). Il est ainsi possible de faire des erreurs lors d'un exercice de lecture qu'on ne ferait pas forcément en parole spontanée.

Par exemple, la plupart des étudiants savent qu'on prononce « tout de suite » /tut sɥit/, mais s'ils voient le mot écrit, ils risquent de le prononcer tel qu'il est écrit, prononcer le schwa et éliminer ainsi le contexte d'assimilation, ce qui serait alors plus éloigné de la forme employée en parole spontanée par les francophones (bien qu'acceptable dans le sud de la France) et par le locuteur lui-même. Mais si certains comportements dans des structures similaires produisent des résultats similaires, et ce, de manière récurrente, on peut dire qu'il ne les maîtrise pas.

On peut tout de même supposer que dans la plupart des cas, si l'apprenant maîtrise parfaitement une règle, il ne devrait faire d'erreurs ni dans la lecture, ni dans la parole spontanée, sauf erreurs exceptionnelles (distraction, fatigue en fin de texte, etc.). De même, si l'apprenant ne maîtrise pas du tout une règle phonétique, il devrait faire des erreurs dans la lecture, alors qu'en parole spontanée, soit il en ferait aussi, soit il éviterait complètement les structures inconnues.

Enfin, d'après plusieurs études, notamment celles de Pagliano et Laks (2008), c'est dans la lecture qu'il y a le plus de liaisons. Pour pouvoir déterminer les compétences des étudiants, il vaudra mieux déterminer les contextes qui favorisent l'apparition de certains phénomènes.

Dans l'ensemble, la lecture comprend plus d'avantages que les autres méthodes. Pour notre recherche sur les liaisons, c'est elle qui reflète le mieux les compétences des sujets. On peut évaluer avec plus de précision les contextes que l'on veut tester et on peut mieux évaluer

les compétences réelles des apprenants. C'est pourquoi ce type d'exercice a été choisi plutôt que les trois autres.

Les études précédemment citées ont utilisé la parole spontanée, la parole semi-guidée et la lecture (pour deux d'entre elles) et ont obtenu des résultats similaires dans les trois sortes de tests. La raison en est sans doute que dans la parole spontanée et semi-guidée comme dans la lecture, n'étaient présents que les contextes les plus fréquents, donnant ainsi les mêmes résultats, aux pourcentages élevés de liaisons réalisées.

## **2.5. Comparaison de groupes différents**

Plusieurs études comparent les productions d'étudiants différents, par exemple les productions d'étudiants qui sont partis en France et ceux qui sont restés chez eux (Howard 2004 et Thomas, 2002) ou bien les étudiants avant et après un séjour en France (Thomas 2002, 2004, 2008). De telles comparaisons fournissent des éléments importants permettant de tirer des conclusions générales sur les éléments maîtrisés en moyenne par niveau, les différences entre les groupes, et les progrès effectués après un séjour linguistique. Toutefois, si l'on peut remarquer des progrès, il est difficile d'établir avec certitude une relation de causalité entre les progrès effectués et le contexte d'apprentissage.

Il est possible que les étudiants qui décident de partir en France soient moins avancés que ceux qui restent au pays, et qu'ils aient ainsi décidé que c'est pour eux une façon d'améliorer leur français. C'est ce que semble indiquer par exemple l'étude de Thomas (2002) où le groupe qui allait partir en France a produit 86.7% de liaisons obligatoires, alors que ceux qui restaient en Ontario en ont produit 90.8%. Il est possible qu'au contraire, seuls les plus motivés, ou les plus auditifs ou les plus kinesthésiques qui apprennent le mieux en parlant, par l'imitation, profitent d'un tel séjour.

Si l'on compare les mêmes étudiants, d'une année sur l'autre, il est difficile de déterminer les facteurs qui sont le plus directement liés aux progrès. En une année, beaucoup de choses se passent. Il arrive qu'on enseigne des règles de grammaire ou de phonétique, qu'on les répète, que les progrès des étudiants soient limités et n'apparaissent que l'année d'après. Puis les étudiants se mettent à maîtriser une règle, à utiliser correctement des formes qu'ils n'arrivaient pas à utiliser correctement jusqu'à présent (apprentissage en forme de cloche).

Cela expliquerait d'ailleurs pourquoi les sujets de l'étude de Thomas (2002) étant restés en Ontario, ont eux aussi fait des progrès avec les liaisons. Il est possible qu'il y ait une certaine « maturation » et que la fréquence et les répétitions participent activement à l'acquisition. De nombreux facteurs peuvent influencer les résultats, tels que la motivation, le temps passé à écouter ou à parler la langue, les relations avec les natifs, etc. Même si l'on remarque des progrès pour le groupe parti en France, et qu'on remarque également des progrès pour le groupe témoin, il est difficile d'en établir les causes exactes.

Il faut aussi considérer avec précaution les comparaisons entre des groupes de 2<sup>e</sup> année et des groupes de 3<sup>e</sup> année qui ne sont pas les mêmes étudiants d'une année sur l'autre. À la fin d'une année scolaire, certains étudiants plus faibles ne continueront pas leurs études en français. D'une part, il y a des pré-requis notamment au niveau des notes obtenues chaque année (dans certaines universités et certains programmes) pour continuer dans un programme de français; d'autre part, si les étudiants ont eu trop de difficultés, ils comprennent qu'il vaut mieux choisir un autre domaine dans lequel ils réussiront mieux. Enfin, les meilleurs étudiants francophones ou ayant fait leur scolarité dans des écoles francophones sont souvent placés directement en 3<sup>e</sup> année (par un test de placement ou avec permission du département). Ainsi, en 3<sup>e</sup> année, il y a un plus petit pourcentage d'étudiants faibles qu'en 2<sup>e</sup> année et un plus grand pourcentage d'étudiants forts.

C'est pourquoi il nous semble préférable de comparer les productions des mêmes sujets d'un même niveau, placés dans les mêmes conditions, ou plus précisément, les conditions les plus proches possibles (il y a toujours des facteurs individuels difficiles à contrôler). Dans ce cas, si l'on compare plusieurs groupes en changeant seulement un facteur, on pourra un peu mieux contrôler les conditions et donc tirer des conclusions un peu plus fiables.

## **2.6. Comparaison de productions de francophones et comparaisons de tests**

Comparer les productions spontanées d'anglophones et de francophones, et les productions de diverses études, permet de fournir des informations générales. Il faut tout de même évaluer les limites de telles comparaisons. Même en se restreignant aux liaisons obligatoires, les méthodes utilisées et les éléments testés peuvent créer des résultats assez différents.

Tout d'abord, les catégories « obligatoires » et « facultatives » ne sont pas uniformément définies. Par exemple l'auxiliaire peut être considéré comme étant un contexte de liaison facultative pour certains (Abry et Chalaron 1994 : 112 ; Thomas, 2002) et obligatoire pour d'autres (Léon, 1966: 120 ; Mastromonaco, 2000). L'Académie Française<sup>34</sup> stipule que la liaison est obligatoire « entre l'adverbe et le mot unis étroitement : trop (p) étroit; bien (n) aise; » alors que Léon (1966 : 129) considère non seulement que c'est une liaison facultative, mais que c'est même une liaison facultative « peu fréquente ». Bref, d'une étude à l'autre, les termes *liaisons obligatoires* et *liaisons facultatives* varient.

Ensuite, si on a un nombre restreint de locuteurs natifs, des productions idiosyncratiques peuvent partiellement modifier les résultats. Ou bien, si les productions sont plus longues et les sujets de conversation plus variés, les contextes de liaison peuvent

---

<sup>34</sup> <http://www.academie-francaise.fr/langue/questions.html> p.18

changer; les étudiants peuvent utiliser des contextes moins fréquents (et donc plus difficiles pour eux et peut-être pas encore maîtrisés) et ainsi donner des résultats différents. Cela pourrait d'ailleurs expliquer les résultats plus faibles des productions des sujets de Howard (2005) lors d'un entretien d'une heure (82.2% de liaisons obligatoires produites par les étudiants avant de partir en France) que ceux de Mastro Monaco (2000) et Thomas (2002).

S'il s'agit d'études en parole spontanée, les conditions d'élocution et de nombreux facteurs rentrent en jeu pour les productions de liaisons facultatives, tels que le contexte phonétique (certaines consonnes sont plus sujettes aux liaisons que d'autres) ou la longueur des mots avant ou après la liaison. Par exemple la liaison est plus fréquente si le deuxième mot est plus long que le premier (Howard, 2004 : 149).

Les interviews, émissions téléphoniques et autres observations sur les francophones fournissent des informations précieuses sur les comportements linguistiques de certains groupes de personnes dans un contexte donné. Par exemple, les études de Léon et Tennant (1990) sur les productions de liaisons des écrivains dans l'émission « Apostrophes » de Bernard Pivot (qui s'est déroulée en France de 1975 à 1985) ou du français utilisé dans les médias de Howard (2006) reflètent certaines tendances des personnes d'un certain âge, dans un contexte formel, dans une situation de prestige et de pouvoir (des milliers de personnes vont regarder l'émission). De même, pour l'étude des politiciens d'Encrevé (1988), qui reflète la tendance de production de liaisons de politiciens français, de haut niveau socio-économique. Par exemple, ces politiciens produisent entre 18 et 67% de liaisons facultatives, avec un peu plus de la moitié d'entre eux (11 sur 21) qui produisent plus de 47% de liaisons facultatives (entre 47% et 67%).

Les résultats des études sur les liaisons facultatives dépendent de plusieurs facteurs, et comme ces facteurs diffèrent d'une étude à l'autre, il est assez difficile de les comparer. Par exemple, on peut comparer les liaisons dans certains contextes, mais il sera difficile de tirer

des conclusions certaines quand les mêmes personnes réalisent la liaison dans un cas et ne la réalisent pas dans un autre, dans exactement le même discours et le même contexte syntaxique.

Il est également ambigu de comparer un test de lecture dans lequel on teste des éléments plus précis, avec la parole spontanée, où certains éléments étudiés sont absents. On peut alors considérer différentes études, plutôt à titre informatif et remarquer des tendances générales communes.

Par exemple, alors que les francophones de Malécot (1975) produisent 96.9% de liaisons obligatoires, ceux de l'étude de Booij et De Jong (1987) ne produisent que 61.2% de liaisons obligatoires et 21.7% de liaisons facultatives. L'écart de plus de 30% sur ce qui est censé être « obligatoire » montre les problèmes méthodologiques liés aux comparaisons d'études différentes en parole spontanée.

Si l'on considère certains contextes de l'étude de Booij et De Jong (1987 : 1010) qui appartiennent pourtant à la catégorie [dét. + nom], tels que *grandes averses*, *anciennes amies*, *certaines exécutants*, *autres activités*, on comprend pourquoi les pourcentages sont plus bas que dans d'autres études. En effet, ce contexte [adj. (qualificatifs et indéfinis) + noms] est beaucoup plus rare que d'autres déterminants et il est parfois considéré comme étant une liaison facultative.

Comme on ne peut jamais comparer les mêmes mots en parole spontanée puisque les locuteurs (et en particulier les locuteurs francophones) choisissent de parler de ce qu'ils veulent et utilisent les mots de leur choix, on ne peut considérer que des catégories syntaxiques (ici [dét. + nom]). Or, à l'intérieur des mêmes catégories, il existe des sous-catégories qui peuvent produire des résultats très différents. C'est pourquoi considérer une catégorie générale de type [dét. + nom] est particulièrement problématique parce qu'on ne compare en fait pas forcément la même chose.

Pour vraiment comparer la même chose, il faudrait au minimum considérer des éléments de type [article défini pluriel + nom de même fréquence et de même longueur] ayant la même consonne de liaison (et le même contexte phonétique). Si l'on avait ces mêmes contextes réunis et qu'on demandait à plusieurs groupes francophones de créer des phrases avec ces mots ou de lire des phrases contenant ces groupes de mots, on trouverait sans doute des résultats avoisinant ceux de Booij et De Jong (1987).

En effet, dans la production de liaisons, le contexte syntaxique ne constitue qu'un des facteurs participant à la liaison. Il y en a d'autres tout aussi importants : les facteurs phonétiques et fréquentiels. Comme en parole spontanée, le cumul des facteurs est forcément différent. Si l'on ne prend en compte qu'un seul facteur dans une catégorie syntaxique, on peut considérer que l'analyse est incomplète.

Si l'on compare les études sur la production de parole spontanée des apprenants L2 de Thomas (2002), Howard (2005) et Mastromonaco (2000), avec les francophones de l'étude de De Jong (1989), on pourrait en déduire que les anglophones de ces études font plus de liaisons obligatoires que les francophones. Mais si l'on compare les liaisons facultatives des apprenants L2 de Howard qui en ont produit 12.7% et ceux de Mastromonaco (2000) qui en ont produit 14% aux jeunes des banlieues de Laks (1983) qui n'en ont produit que 4%, on pourrait en déduire que les anglophones en produisent trois fois plus que les francophones.

Il est possible qu'ayant reçu un modèle académique favorisant l'utilisation du registre recherché (qui encourage la réalisation de liaisons facultatives), ils produisent de manière systématique certaines liaisons facultatives avec *c'est* ou *pas*, et puissent en réaliser certaines en plus grande proportion que les francophones parlant un français familier ou populaire.

Cela serait théoriquement possible si les apprenants L2 maîtrisaient parfaitement le français, les registres de langue et la variation sociolinguistique. Cela ne fait aucun doute que de nombreux linguistes, professeurs d'université et autres spécialistes du français L2, peuvent

maîtriser la langue écrite et/ou orale à un niveau supérieur à la moyenne des francophones et puisse produire plus de liaisons (facultatives) qu'eux.

Or, si l'on considère les productions de ces apprenants, on voit que ce n'est pas le cas, leur maîtrise du français est limitée. Il est donc peu probable que ces étudiants L2 produisent trois fois plus de liaisons facultatives que des francophones. On trouvera sans doute que les productions de francophones contiennent des catégories et sous-catégories plus variées, des contextes plus rares, ce qui rend alors la comparaison difficile.

Comme les apprenants L2 utilisent les mots les plus fréquents, et des structures moins complexes, ils font aussi les liaisons les plus fréquentes qui correspondent à ces contextes, ce qu'on voit par exemple dans les productions des sujets de Matromonaco (2000). Par contre, les francophones utilisent des contextes plus rares qui sont en fait presque des liaisons facultatives, ce qu'on voit dans les études de francophones telles que celles de Ågren (1973), Malécot (1975), Booij et De Jong (1987), Gadet (1989), Howard (2006) ou Mallet (2008).

Certes, les francophones utilisent eux-aussi les structures les plus fréquentes de type [dét. + nom], mais ils utilisent dans les mêmes productions des structures un peu moins fréquentes quasi absentes (ou complètement absentes) des productions spontanées des anglophones, notamment les contextes de liaisons interdites. De plus, ils utilisent des items moins fréquents et des mots plus longs.

Étant donné que la séparation entre liaisons obligatoires et facultatives est floue, dans un même contexte tel que [dét.+ nom] ou [prép. + GN], la comparaison s'avère difficile, si l'on ne compare pas les mêmes prépositions et les mêmes déterminants.

Il faut en effet considérer que le nombre de contextes possibles est plus faible chez les non-natifs. Dans l'étude de Hieke (1984), par exemple, on voit que, si on enlève les pauses (qui empêchent des contextes possibles), les anglophones (natifs) ont produit une moyenne (pour 100 syllabes) de 15.60 contextes d'enchaînements possibles et 12.11 cas



d'enchaînements réels, soit 77.63% d'enchaînements réalisés parmi les contextes possibles. Les Allemands (apprenant l'anglais) ont produit 12.86 (pour 100 syllabes) de contextes d'enchaînements possibles et en ont réalisé 6.88 soit 53.50% parmi les contextes possibles.

Ceci représente déjà des différences importantes. Mais si l'on considère les contextes de départ sans tenir compte de ce qui a empêché les liaisons et enchaînements (pauses, hésitations et reprises), le nombre plus restreint de contextes liants et le pourcentage final d'enchaînements effectués, on arrive en réalité à un nombre très inférieur à celui des locuteurs natifs. En fait, les non-natifs ont produit 6.88 enchaînements, tandis que les natifs en ont produit 12.11, soit pratiquement le double, pour 100 syllabes.

Pour pouvoir réellement comparer la « maîtrise » d'un certain point, d'un groupe par rapport à un autre, il faudrait comparer la même chose. Si l'on compare la production de [être + verbes au participe passé] ou [en + nom], on peut relativement bien comparer les contextes. C'est ce qu'a fait Thomas (2002) en comparant les liaisons avec les différentes personnes d'*être* (*suis, es, est, etc.*). Toutefois, parler de « liaisons obligatoires », « facultatives », ou « interdites » ou même parler de catégorie [dét. + nom] ou [prép. + GN] est trop vague tant qu'on n'est pas sûr que les points testés soient exactement les mêmes (même préposition, avec un mot<sup>2</sup> de même longueur, de même fréquence et de même fréquence de cooccurrence).

Comme le remarquent Larsen-Freeman et Long (1991), les méthodologies employées pour les recherches en acquisition d'une L2 produisent des résultats différents selon qu'on utilise une méthode qualitative/quantitative, longitudinale/transversale, selon qu'on utilise des exercices de lecture à haute voix, des exercices de structure, d'imitation, de traduction, des exercices guidés ou semi-guidés. Enfin, l'interprétation des données peut être très différente selon les chercheurs, selon le cadre théorique de la recherche, et selon d'autres paramètres.

De plus, les critères pris en considération pour les tests ont des implications directes sur les résultats. Par exemple, si l'on considère qu'une liaison est réalisée, même si elle n'est

pas enchaînée, ce qui est le cas dans les études de Thomas (2002) et de Mastromonaco (2000), on obtient des taux de liaison plus élevés que si l'on considère, comme dans notre étude, que la liaison n'a pas été réalisée (correctement) si elle n'a pas été enchaînée. De même, Mastromonaco (2000) ne prend pas en compte si ce sont les mêmes mots ou expressions qui ont été répétés en parole spontanée, et considère seulement les cas réalisés ou non-réalisés.

Selon Green et Hinze (1990), les résultats sont différents selon les études pour les raisons suivantes :

The researchers are using different definitions of basic categories and are collecting data for rather different purposes. The second (reason) has to do with the interaction of variables: it is well known that the higher the number of variables in an experiment the harder it is to control for their effects. [...] Discrepancies in the results of empirical studies can usually be traced to one of three factors: differences in the subcategorization of linkage types; and differences in the computation of the ratio of actual to potential incidence. The factors are not mutually exclusive and can indeed, be mutually compounding. (Green et Hinze, 1990 : 75-76),

C'est exactement le problème qui se pose dans les études en question. Si l'on compare deux groupes en parole spontanée, et qu'on ne considère que le facteur syntaxique, l'influence des autres facteurs n'est pas considérée, modifiant alors l'interprétation des données.

C'est la raison pour laquelle nous pensons que la meilleure façon de comparer des groupes différents d'étudiants avec des francophones est, quand c'est possible, de comparer leur production dans exactement le même contexte, et le seul contexte qui puisse être le même est la lecture à haute voix du même texte. Uniquement dans ce cas, peut-on comparer la même chose, les mêmes items, dans les mêmes contextes. C'est pourquoi nous avons choisi de comparer la lecture d'anglophones et de francophones lisant à haute voix le même texte.

## 2.7. Les comparaisons de lectures de plusieurs études

Il est possible de comparer les lectures de plusieurs études, si l'on compare des contextes similaires. On peut alors comparer les contextes [dét. pluriel + nom], [aux. être + part. passé] etc., si les mots sont de fréquence comparable, de longueur comparable, si les contextes sont vraiment semblables.

Toutefois, si l'on considère seulement une même catégorie syntaxique, telle que le contexte [dét. + nom], les contextes peuvent en réalité être assez différents. Par exemple, dans le texte 1 de Mastromonaco (2000 : 299), on trouve dans la catégorie [dét. + nom] : un adjectif qualificatif + nom (*plein air*), un adjectif possessif + nom (*mon avis*), un article indéfini singulier + nom (*un arrangement*) et trois articles indéfinis pluriel + nom, (*des ingrédients, des emballages, des aliments*).

Parmi ces six déterminants, se trouvent quatre articles les plus fréquents : *un, des, des, des*. *Mon avis* est pratiquement lexicalisé, car la plupart des étudiants savent dire *à mon avis* en faisant la liaison. En effet, on retrouve cette séquence souvent dans les productions des sujets de Mastromonaco (2000 : 299-348). Il ne reste comme unique difficulté que *plein air* qui est d'ailleurs une expression presque figée.

Si l'on compare les deux textes utilisés par Mastromonaco : le texte 1 lu par les sujets 1-19 (p. 299-320) et le texte 2 lu par les sujets 20-59 (p. 320-348), les deux textes contiennent des difficultés différentes et même différents niveaux de difficultés. Cela explique en partie, pourquoi, par exemple, pour la lecture des deux textes il y eu des différences statistiques significatives dans la production de liaisons facultatives. Nous voyons ici les difficultés de comparer des textes apparemment de difficulté similaire.

Si un autre texte examine par exemple 40 cas de liaisons obligatoires en contexte [dét. + nom] avec différentes sous-catégories (h aspiré, semi-consonne, nasale, etc.), avec des mots rares (ce qui est le cas avec les mots commençant par un h aspiré), il y aura sans doute plus de

chances pour que les étudiants rencontrent des contextes moins familiers et fassent donc moins de liaisons. C'est la raison pour laquelle, nous avons choisi de comparer des groupes de francophones et d'anglophones lisant le même texte dans les mêmes conditions, et la méthode adoptée pour faire cette comparaison sera exposée dans la section qui suit.

### 3. La procédure

Nous allons donner ici un court résumé de la procédure utilisée. Une description plus détaillée de chacun des éléments présentés sera donnée dans les prochaines sections.

Nous ne pouvions pas modifier l'emploi du temps prévu dans le cours de laboratoire. Notre présentation ne pouvait se faire que dans un temps restreint : nous n'avions qu'un total de 110 mn, étalé sur trois semaines.

Le tableau suivant résume ce qui s'est passé dans chaque session pour tous les groupes :

	<b>SESSION 1</b>	<b>SESSION 2</b>	<b>SESSION 3</b>
contenu	Brève explication du projet <b>pré-test</b>  Questionnaire (1)	entraînement phonétique (cours sur les liaisons)  <b>test</b>	<b>post-test</b>  Questionnaire (2)
temps	40 mn	40 mn	30 mn

**Tableau 3.1. Les sessions de l'étude.**

**Lors de la session 1**, nous avons expliqué brièvement aux étudiants le but de l'étude et la façon dont les événements allaient se dérouler. Les étudiants ont rempli le formulaire de consentement qui précisait qu'ils pouvaient se retirer de l'étude à tout moment.

Puis, les étudiants ont répondu au questionnaire 1 portant sur leur formation linguistique, le nombre d'années ayant étudié le français, les activités en français, etc. (voir le questionnaire en annexe 3). Ils ont fait un pré-test qui consistait à s'enregistrer sur CAN8<sup>35</sup> en

<sup>35</sup> On peut trouver des renseignements complémentaires sur le site <http://www.can8.com/>

lisant le texte portant sur les liaisons et les enchaînements, que la compagnie Pearson (qui produit Can8) décrit de la façon suivante :

The CAN-8 VirtuaLab is a 100% digital language laboratory. It is a network based, interactive, multimedia learning facility with the ability to teach and test languages; from Arabic to English to Chinese. With voice over the Internet (VOIP) since 1997, this comprehensive learning system is used for classroom or distance learning, written or spoken language testing and group conferencing.

Les étudiants avaient déjà utilisé CAN8 dans le cours de laboratoire, il n'était donc pas nécessaire de leur expliquer comment s'en servir. Un fichier étant réservé à l'enregistrement, ils n'avaient qu'à s'enregistrer dans le fichier qui leur avait été indiqué.

**Lors de la session 2**, une semaine plus tard, les étudiants ont reçu une courte formation sur les liaisons et les enchaînements. Dans toutes les sections, nous avons fait une explication formelle du principe de syllabation ouverte en français et de la contrainte anti-hiatus ; nous avons donné des exemples de chaque catégorie (liaisons obligatoires, interdites, enchaînements) ; nous avons fait répéter aux étudiants plusieurs séquences dans chaque catégorie ; nous avons fait des commentaires sur les erreurs entendues et nous avons fait répéter plusieurs fois les séquences ayant créé des difficultés. Par contre, le moment de l'enregistrement était différent selon les sections, afin de déterminer quelle partie du cours produisait quel effet. Pendant cette session, les étudiants de tous les groupes ont fait un test de lecture en s'enregistrant sur CAN8 en lisant toujours le même texte.

**Lors de la session 3**, une semaine après la mini-formation, les étudiants ont refait le même test d'écoute que lors des première et deuxième sessions ; ils ont refait le même enregistrement et ont répondu à un petit questionnaire (voir le questionnaire en annexe 4) pour savoir comment ils avaient trouvé la formation.

## **4. Les sujets**

### **4.1. Les sujets francophones**

Afin de tester les mêmes éléments dans les mêmes conditions d'énonciation (la lecture à haute voix), il fallait comparer les productions des groupes d'apprenants du français L2 avec les productions de francophones. Toutefois le mot « francophone » peut parfois avoir des sens différents. De nombreuses études portent sur les productions des Français de France (Ågren, 1973 ; Malécot, 1975 ; Ashby 1981, 2003, Pagliano et Laks, 2008, etc.). Il nous paraissait utile de déterminer s'il y avait des différences entre les enseignants francophones vivant en Ontario venant de différents pays et de différentes régions francophones.

Pour cela, il nous semblait pertinent d'analyser les productions de plusieurs groupes de francophones. Chaque groupe dans notre étude est trop restreint pour en tirer des conclusions générales sur les Français, les Québécois, les Franco-Ontariens ou les Africains. La thèse ne porte pas sur les francophones à proprement parler. Nous avons simplement voulu évaluer des tendances générales de plusieurs groupes et sous-groupes, et déterminer les liaisons et enchaînements qui sont réalisés par tous les groupes, par la plupart des groupes et ceux qui sont réalisés en faibles proportions pour tous les groupes. Les francophones servent de groupe témoin.

Les francophones volontaires ont d'abord été « recrutés » au département de français. Nous avons contacté en personne professeurs, assistants et moniteurs de laboratoire. Nous avons pris rendez-vous avec eux, nous nous sommes rendus dans leur bureau. Puis, nous sommes allés dans un centre francophone où nous avons rencontré quatre Franco-ontariennes. Nous leur avons expliqué le but de l'étude et toutes les personnes ont accepté de lire le texte et de répondre au questionnaire (annexe 5).

Tous les francophones ont lu le même texte que les anglophones. Les consignes étaient seulement « lisez le texte de façon ‘naturelle’ ». Aucun francophone n’a demandé de préparer le texte, sauf une seule personne. Une collègue de travail (d’un autre département, n’enseignant pas le français) a tenu à « préparer le texte », sachant qu’il portait sur les liaisons, voulant sans doute produire une bonne impression et « bien prononcer ». Au final, elle n’a pas tenu compte des annotations inscrites dans son texte, et a prononcé de la façon qui lui semblait la plus naturelle quand elle le lisait à voix haute, m’a-t-elle dit après avoir lu le texte, en faisant de nombreuses liaisons qu’elle n’avait pas remarquées en lisant le texte à voix basse, en faisant ses annotations.

Les francophones ont été enregistrés avec le magnétophone digital Panasonic RR/US750 et les enregistrements ont ensuite été téléchargés sur ordinateur, pour conserver les fichiers sonores et les classer avec différents titres et catégories. Comme les anglophones, les francophones ont lu et signé le protocole d’accord autorisé par le comité éthique.

Le groupe francophone se compose de 20 sujets issus de diverses origines géographiques. Nous avons choisi de prendre comme groupe témoin des francophones de diverses origines, afin de déterminer si la production de différents groupes et sous-groupes se rapproche des normes orthoépiques et des résultats des études précédentes sur les productions de liaisons de francophones, pour pouvoir ainsi déterminer les tendances générales de tous les groupes francophones ainsi que les principales différences dans la lecture du même texte.

Nous avons aussi pour but d’évaluer les effets de fréquence sur différents groupes francophones ayant appris le français dans des conditions différentes. Il est ainsi possible de tester le rôle de la fréquence des structures syntaxiques, le rôle de la fréquence lexicale, ainsi que le rôle des consonnes de liaison pour les groupes francophones majoritaires et les groupes francophones minoritaires.

Voici une description des sujets et des critères utilisés pour déterminer à quel groupe ils appartiennent, à partir des questions du questionnaire et en discutant avec eux.

Les francophones majoritaires (Français, Belges, Québécois et bilingues) correspondent aux critères suivants :

1. Personnes ayant parlé le français à la maison pendant l'enfance, et/ou
2. Personnes ayant au moins un parent francophone qui leur a toujours parlé en français pendant leur enfance, et/ou
3. Personnes ayant été dans un milieu francophone majoritaire, y compris, mais pas seulement, à l'école pendant leur jeunesse.

Les francophones « bilingues » (qui appartiennent au groupe de francophones majoritaires) correspondent aux critères suivants :

1. Personnes ayant appris deux langues en même temps, par exemple si un parent était francophone et l'autre anglophone dès la naissance, et/ou
2. Personnes ayant appris les deux langues (en même temps ou l'une après l'autre) très tôt dans l'enfance (avant l'âge de 10 ans), et/ou
3. Personnes qui se définissent comme « bilingues », comme ayant deux langues maternelles.

Les francophones « minoritaires »<sup>36</sup> correspondent aux critères suivants :

1. Personnes ayant parlé le français seulement à la maison et avec un groupe de locuteurs restreints : amis de la famille, famille, (mais pas à l'école, ni ailleurs), et/ou
2. Personnes ayant parlé le français seulement à l'école, mais pas à la maison, et pas ou rarement en dehors de l'école (langue maternelle africaine), et/ou

---

<sup>36</sup> D'après le dictionnaire de linguistique et sciences du langage (2007 : 304), les définitions de langue minoritaire sont les suivantes : Une langue est dite *minoritaire* quand, dans une étendue donnée, elle est moins parlée qu'une autre, dite majoritaire. Une langue est dite également *minoritaire* quand elle est le fait d'une minorité nationale, c'est-à-dire d'une communauté qui n'a pas pu faire triompher son droit à l'indépendance ou au moins à l'autonomie culturelle.



3. Personnes ayant été pendant leur acquisition du français dans une situation où le français
4. était moins parlé que la langue dominante de la région (ou du pays).

Nous avons défini comme « francophone » toutes les personnes ayant appris le français avant « la période critique » d'environ 10-12 ans (pour un résumé des études portant sur la période critique se référer par exemple à Birdsong, 2003, et Bongaerts, 2003). La limite entre francophone majoritaire, francophone minoritaire ou francophone bilingue est parfois très fine.

Au Canada, de nombreuses personnes parlent les deux langues nationales, le français et l'anglais, et à ce titre elles pourraient être considérées comme « bilingues ». Par exemple, les sujets franco-ontariens de notre étude sont tous bilingues. Les sujets africains de notre étude parlent tous au moins deux langues et la plupart du temps trois langues ou plus et ils pourraient également s'appeler « bilingues » ou « trilingues ». Enfin, certains francophones de notre étude parlent l'anglais à un très haut niveau (en particulier certains sujets québécois) et pourraient aussi être considérés comme bilingues. Ce qui a déterminé la catégorie à laquelle une personne appartient, c'est d'une part si une langue était dominante par rapport à l'autre dans l'enfance, et d'autre part comment la personne se considère face aux deux langues.

Parfois, le français a été la langue maternelle et langue majoritaire les premières années de la vie du locuteur, mais le sujet a déménagé et l'anglais est devenu la langue majoritaire à la maison et à l'école, ou le contraire. Tous les renseignements ne sont pas forcément dans le questionnaire et c'est en parlant avec les sujets que nous avons déterminé si certaines personnes étaient francophones majoritaires et appartenaient à telle ou telle catégorie. Pourtant, la réponse à la question langue maternelle était généralement la réponse la plus importante pour déterminer l'appartenance aux catégories.

Voici les classifications des groupes et des sujets :

► **12 sujets francophones « majoritaires »**

5 sujets européens : 4 sujets français

1 sujet belge

5 sujets québécois

2 sujets bilingues (parents francophones vivant à London, Ontario).

**Les deux francophones bilingues**

Les personnes bilingues ayant été élevées dans au moins deux langues ont été placées dans la catégorie majoritaire et dans la sous-catégorie « bilingue » pour les raisons suivantes. Pour un sujet, la langue maternelle de la mère est le français. La langue maternelle du père est le flamand. Le père a appris le français à l'adolescence. Ses parents lui ont parlé français jusqu'à l'âge de sept ans. Quand la famille est allée vivre en Ontario, ils ont tous commencé à parler anglais, y compris à la maison. Le sujet a parlé français à l'école au Québec jusqu'en deuxième année (à l'âge de 7 ans), puis l'anglais en Ontario après le déménagement.

Pour l'autre sujet bilingue, la langue maternelle de la mère est le français et la langue maternelle du père l'anglais (il est canadien anglophone). Le père lui parlait anglais et la mère lui parlait français. Les parents se parlaient entre eux en anglais. Le sujet répondait dans la langue respective de chacun, c'est-à-dire en anglais au père et en français à la mère. Étant inscrit dans une école francophone, il parlait français à l'école.

Puisque ces personnes ont reçu la plupart de leur input de leur mère française étant petit et que celle-ci leur parlait exclusivement en français, que ces sujets considèrent que le français est leur langue maternelle, ils font partie des francophones majoritaires.

## ► 8 Francophones « minoritaires »

### - 4 Franco-ontariens

Les quatre personnes ont des parents francophones et ont habité dans différentes régions d'Ontario où le français n'était pas la langue majoritaire. Comme ces personnes ont toutes plus de 60 ans, leur enfance remonte à une époque où les francophones n'avaient pas toujours accès à l'éducation en français autant qu'aujourd'hui. Les sujets de l'étude ont seulement parlé le français à la maison et avec un cercle restreint de personnes, et l'anglais à l'école et le reste du temps. L'anglais était pour elles la langue majoritaire.

### - 4 Africains (Sénégal, Burundi, Congo, Mali)

La langue officielle du Sénégal est le français, bien que pour 42% de la population la langue maternelle soit le wolof<sup>37</sup>. Au Burundi, les deux langues officielles sont le français et le kirundi. En République démocratique du Congo, la langue officielle est le français et les langues nationales sont le kikongo, le lingala, le kiswahili et le tshiluba<sup>38</sup>. Au Mali, la langue officielle est le français et il y a 13 langues nationales reconnues.

Les Africains ont appris le français à l'école, leur langue maternelle est une langue africaine et le français est la langue officielle ou l'une des langues officielles de leur pays. Bien que les Africains aient le français comme L2 ou même L3, ils sont unanimement reconnus (par tous les autres francophones) comme étant francophones de par la qualité de leur français. Beaucoup d'entre eux sont professeurs de langue dans les universités et enseignent dans les écoles canadiennes et servent de modèle linguistique aux apprenants. Enfin, nous souhaitons qu'ils participent à l'étude parce qu'ils avaient eu des conditions d'apprentissage différentes des autres francophones.

Voici quelques renseignements supplémentaires sur les sujets francophones quant à leur âge, leur sexe, et leur origine géographique dans les tableaux 3.2. et 3.3.

---

<sup>37</sup> <http://www.tlfq.ulaval.ca/axl/afrique/senegal.htm>

<sup>38</sup> <http://www.tlfq.ulaval.ca/axl/afrique/czaire.htm>

	Europe					Canada					Bil.	
FR.MAJ	1	2	3	4	5	6	7	8	9	10	11	12
1. Sexe	F	F	F	M	F	F	M	F	M	M	M	M
2. Age	28	23	22	28	45	44	30	39	25	34	36	49
3. Pays de naissance	France	France	France	France	Belgique	Canada	Canada	Canada	Canada	Canada	C. (bil)	C. (bil)

**Tableau 3.2. Renseignements du questionnaire des sujets francophones (majoritaires)**

	Franco-ontariens				Africains			
FRANC. MIN.	1	2	3	4	5	6	7	8
1. Sexe	F	F	F	F	M	M	F	F
2. Age	65	61	71	61	30	29	26	23
3. Pays de naissance	Canada	Canada	Canada	Canada	Congo	Sénégal	Burundi	Mali

**Tableau 3.3. Renseignements du questionnaire des sujets francophones (minoritaires)**

La moyenne d'âge des locuteurs francophones est de 38.45 ans. Tous les sujets ont fait au moins trois ans d'études universitaires, sauf une personne franco-ontarienne qui a obtenu un diplôme après deux ans au « collège » (l'équivalent français d'un BTS). Il y a huit hommes et douze femmes.

D'après le MBU, il semblerait logique que les francophones minoritaires et les francophones « bilingues » ayant appris les deux langues en même temps, avec l'un des parents bilingues, subissent l'influence de l'autre langue, surtout si l'apprentissage de ces deux langues s'est fait pendant l'enfance. Ces personnes ont reçu un grand input dans une autre langue que le français, ce qui veut dire que le contact avec le français a été plus restreint que pour un monolingue qui aurait reçu 100% de son input dans une seule langue.

Or, les conditions d'apprentissage sont cruciales dans l'acquisition des liaisons, comme on le voit dans l'acquisition d'une L1 chez les enfants, dans l'étude de Chevrot, Dugua et al. (2005). Ce qui influence une L1 c'est d'une part combien la L1 est parlée et d'autre part combien est parlée la L2 (ou la L3). Si la L2 est la langue majoritaire, il y a de fortes chances pour que cette situation de contact de langues influence la L1 (voir les travaux de Beniak et Mougeon (1989) sur le français en Ontario).

D'après certaines études, chaque langue parlée peut influencer les autres. Par exemple, la L3 des élèves de l'étude de Griessler (2001) a une influence sur la L2. On voit aussi l'influence de l'anglais sur le français des Turcs ayant déjà appris l'anglais qui apprennent le français dans l'étude de Özçelik (2008).

Si la L2 a été parlée dans l'enfance autant que la L1 et finit par être parlée plus que la L1, la L2 devient la langue dominante. Choi (2003 :108) fait remarquer que la langue « forte » influence la langue « faible ». De même, Belletti et Hamann (2000) remarquent un transfert chez certains sujets (mais pas tous) dans l'acquisition bilingue d'une langue sur une autre, dépendant de plusieurs facteurs. La limite entre acquisition bilingue et acquisition d'une L1 et d'une L2 devient alors très floue. Nous souhaitons ainsi aborder les différences dans les conditions d'apprentissage pour déterminer les différences de production entre les groupes et pour observer des tendances communes à plusieurs groupes.

De nombreux facteurs devraient influencer sur la qualité du français des francophones : la façon dont la langue a été apprise, l'existence d'une ou plusieurs langues maternelles, une autre langue parlée régulièrement ou apprise dans l'enfance, ou une autre langue parlée à la maison, la langue dominante quotidienne, etc.

## **4.2. Les sujets anglophones**

Nous avons contacté le directeur du cours de 1<sup>e</sup> année de notre université et lui avons expliqué nos projets. Nous lui avons proposé de faire une petite leçon de phonétique dans la partie laboratoire du cours. Il a accepté et a parlé du projet aux assistants de laboratoire qui l'ont eux-mêmes accepté. Ils en ont ensuite parlé à leurs étudiants.

Lors de la première session, nous avons expliqué le but de l'étude et la procédure. Nous avons expliqué aux étudiants qu'ils devaient tous assister à la présentation, tout comme ils devaient assister aux cours de laboratoire. Par contre, ils n'étaient pas obligés de faire les

tests et pouvaient interrompre leur participation aux tests à tout moment. Ils ont rempli des formulaires et signé la feuille de protocole du comité éthique.

Tous les étudiants sauf 2 ont choisi de participer à l'étude et donc de faire les enregistrements. Sur 55 sujets qui ont participé à l'étude au départ, certains ont manqué l'une des trois sessions, d'autres n'ont pas rempli le questionnaire. Certains sont arrivés en retard, d'autres ont eu des problèmes avec les enregistrements ou d'autres problèmes techniques. Au final, il restait 37 sujets qui avaient rempli les questionnaires, fait les trois enregistrements et dont les enregistrements étaient audibles.

Voici ici quelques informations concernant les sujets anglophones qui ont choisi de participer à l'étude. Nous appelons l'ensemble de tous les groupes d'apprenants, le groupe « anglophone », même si plusieurs sujets n'ont pas l'anglais comme langue maternelle. Toutefois tous les sujets parlent l'anglais et la majorité est anglophone.

Sur les 37 sujets, il y avait:

- 27 anglophones

- |   |                              |
|---|------------------------------|
| - 7 bilingues (nés au Canada) parlant l'anglais et  | - le serbe                   |
|   | - le créole                  |
|   | - le punjabi                 |
|   | - le cantonais               |
|   | - le mandarin                |
|   | - le suédois                 |
|   | - le kashi (langue indienne) |
| - 1 sujet parlant                                   | - le cantonais               |
| - 2 sujets parlant une autre langue que l'anglais : | - l'allemand                 |
|   | - l'arabe                    |

Nous avons séparé le locuteur du cantonais car cela pouvait être utile pour comparer l'effet de la langue maternelle, et le cantonais est une langue très différente du français

Avant de décrire le parcours scolaire des sujets, expliquons brièvement le système scolaire du Canada en ce qui concerne l'apprentissage du français langue seconde. Pour apprendre le français langue seconde, voici les programmes offerts en Ontario (et dans le reste du Canada) tels que le décrit le site du gouvernement ontarien<sup>39</sup>.

*Core French* (programme de base de français) – Les élèves suivent un cours de français. Ils acquièrent une maîtrise de base de la langue.

*Extended French* (programme intensif de français) – Les élèves suivent un cours de français et au moins une matière du curriculum est enseignée en français, les mathématiques par exemple. Au moins 25 pour 100 de l'enseignement se fait en français.

*French Immersion* (programme d'immersion en français) – Les élèves suivent un cours de français et au moins trois matières du curriculum sont enseignées en français, les mathématiques par exemple. Au moins 50 pour 100 de l'enseignement se fait en français.

Le programme intensif de français se rapproche du principe du français de base dans la mesure où l'emphase est sur l'étude de la langue tout en recevant entre trois et quatre fois le nombre d'heures d'instruction du français de base.

Précisons également les principes du système francophone. Celui-ci était à l'origine réservé aux enfants francophones dont les parents étaient francophones. De nos jours, au moins en Ontario, pour entrer en école francophone à l'âge de 4 ou 5 ans, il n'est pas toujours nécessaire que les parents des élèves soient francophones, du moment qu'ils parlent le français et répondent à certains critères déterminés par l'école (il faut généralement qu'au moins un parent puisse parler le français à un niveau suffisant pour pouvoir aider son enfant avec les devoirs, les lectures et autres activités en français). Les conditions d'accès aux non francophones peuvent varier d'une école, d'une région et d'une province à l'autre, la priorité étant évidemment donnée aux francophones. Lorsque les parents ont rempli un questionnaire, ont eu un entretien avec un comité évaluateur, et que la direction de l'école considère que les

---

<sup>39</sup> <http://www.edu.gov.on.ca/fre/amenagement/FLS.html>

élèves peuvent suivre leur scolarité en français, certains élèves non francophones peuvent suivre leur scolarité dans le système francophone.

Parmi les 37 sujets anglophones de notre étude,

- 17 sujets ont étudié dans le programme de base (*core French*) ;
- 6 sujets ont étudié le français en école d'immersion ;
- 3 sujets ont étudié le français dans le programme intensif de français (*extended French*) ;
- 4 sujets ont étudié dans le programme de base et dans le programme intensif, (*core French* et *extended French*) ;
- 3 sujets ont étudié dans le programme de base et dans le programme d'immersion (*core French* et *immersion*) ;
- 2 sujets ont étudié dans le programme intensif et dans le programme d'immersion (*extended French* et *immersion*) ;
- 1 sujet a étudié dans le programme de base, dans le programme intensif et dans le programme d'immersion (*core French*, *extended French* et *immersion*).
- 1 sujet a étudié dans le programme de français langue maternelle et dans le programme d'immersion (*francophone* et *immersion*) ;

Aucun des sujets n'a suivi toute sa scolarité dans une école francophone car les étudiants issus de ce système ont généralement un niveau de français comparable ou identique à celui d'un francophone et sont généralement placés au moins en 2<sup>e</sup> année, voire en 3<sup>e</sup> année, tout comme les francophones.

Au moment de l'étude, les sujets étaient tous inscrits dans le même cours (dans différentes sections) de première année de spécialisation en français (niveau intermédiaire A2-B1 d'après le CECR). Les étudiants inscrits dans ce cours ont, pour la plupart, suivi des cours de français jusqu'en 12<sup>e</sup> année à l'école. Sinon, ils ont passé un test de placement qui a déterminé qu'ils devaient aller dans ce cours. Quelques sujets ont fait des séjours linguistiques



en milieu francophone, dont le séjour varie d'un sujet à l'autre, beaucoup n'en ont pas fait et un sujet a passé un an en région francophone. L'âge moyen des 37 étudiants au moment de l'étude était de 19 ans. Il y avait 6 hommes et 31 femmes.

Il y avait quatre groupes, dont le groupe 4 qui était le groupe témoin parmi les anglophones. Malheureusement, le groupe témoin (groupe 4) qui était déjà au départ un petit groupe, s'est progressivement amenuisé puisqu'un étudiant est arrivé en retard, l'autre a manqué un enregistrement, etc. et à la fin, il n'y avait plus que deux sujets dans le groupe.

## **5. Les méthodes d'enseignement utilisées dans l'étude**

Voici d'abord quelques aspects importants qui ont été pris en considération pour décider du choix d'activités de la mini-leçon portant sur l'enseignement de la contrainte anti-hiatus, la syllabation et l'usage des liaisons obligatoires et interdites. Vu les restrictions de temps, les liaisons facultatives n'ont pas été traitées. Il a seulement été précisé que si les liaisons ne sont ni obligatoires, ni interdites elles sont facultatives. Dans ce cas, on a le choix de les faire ou non et plus le registre de langue est soutenu, plus on fait de liaisons.

Voici plusieurs principes importants pris en compte pour l'enseignement de la prononciation. Étant donné le filtre ou « crible phonologique » dont parlait Troubetzkoy (1939) qui empêche les apprenants d'entendre certains éléments, il est essentiel de rendre ces éléments plus saillants et d'attirer l'attention des apprenants sur ce qui est mal perçu ou pas perçu du tout à l'aide d'instruction formelle. Venkatagiri et Levis (2007 : 263) distinguent trois approches pour enseigner certains points qui pousseraient les apprenants à se concentrer sur la forme :

- Le premier type d'instruction est l'approche « non-interventionniste », préconisée par Krashen (1982, 1985). D'après cette approche, les apprenants imitent la L1 y compris la

forme précise. Il n'y a donc pas besoin d'enseignement formel, car d'après cette approche, l'écoute et la répétition suffisent, tout comme les enfants apprennent une langue « naturellement » en l'écoutant et en parlant.

- La seconde approche que Long et Robinson (1998 : 16) appellent « le focus sur la forme » ou en leurs termes : « *focus on formS* » s'apparente à l'approche traditionnelle où des points précis sont enseignés dans un ordre préétabli. Tout est organisé dans un syllabus. Les professeurs enseignent des formes linguistiques grammaticales ou phonétiques, telles que prévues dans le syllabus en se concentrant sur la forme.

- La troisième approche consiste à créer des activités de communication dans lesquelles la précision et la justesse de la forme déterminent la compréhension des locuteurs. L'enseignement de la forme est ici plutôt implicite.

Il existe également de très nombreuses façons d'enseigner la prononciation tout en utilisant une approche « traditionnelle ». Il est possible d'utiliser des chansons ou de la musique (Medina, 2002 ; Salcedo, 2002), des films et cassettes vidéos (Yates, 2003), des ressources sur internet, (Tomé, 2003), des didacticiels (Knoerr, 2000), des cassettes et cédéroms (Knoerr et Weinberg, 2005), la psychodramaturgie (Dufeu, 1986), des cours de phonétique (Abe, 2006 ; Couper, 2003, 2006), des exercices de répétition (Neufeld, 1979, 1980) et bien d'autres.

De nombreux chercheurs s'accordent sur les effets positifs de l'instruction formelle et de la correction tels que Lightbown (2003) et Spada (2007). La vaste majorité des études portant sur le rôle de l'enseignement de la prononciation indiquent que le groupe expérimental ayant reçu une instruction formelle fait de nets progrès, alors que le groupe témoin n'ayant pas reçu d'instruction, soit n'en fait pas, soit en fait moins que le groupe expérimental. C'était le cas par exemple dans les études de Derwing et Rossiter (2003), Blanche (2004), Boubnova (2000), Dufeu (2008), Syrovatskaja (2008) et Yates (2003).

S'il peut paraître un peu surprenant que les étudiants fassent des progrès significatifs quelle que soit la méthode utilisée (dictée, répétitions, exercices de discrimination, cours de phonétique, utilisation de didacticiel, etc.), toutes ces méthodes ont plusieurs points en commun dont le fait qu'elles poussent les apprenants à faire attention à la forme en se concentrant sur les différences entre leur propre production (leur interlangue) et la forme cible.

Schmidt (1990) considère que le plus important dans le développement d'une L2 est la « prise de conscience » (ou en ses propres termes « *noticing* ») de certains phénomènes ou certaines formes linguistiques. Dans des cours de prononciation ou phonétique ayant cet objectif, l'apprenant se concentre sur la forme, écoute plusieurs fois des mots, séquences et phrases et reproduit les formes cibles en ayant de préférence du feedback sur sa production ou toute autre forme de correction.

Pour que les apprenants fassent attention à la forme, ils ont besoin que cette forme soit mise en relief, portée à leur attention, mise à leur disposition. Ils doivent entendre la forme correcte plusieurs fois, bien l'entendre, être capables de discriminer certaines formes, comprendre certains phénomènes, pratiquer certaines formes jusqu'à ce qu'elles deviennent de plus en plus familières et idéalement jusqu'à ce qu'elles deviennent naturelles et automatiques.

D'après les études précédemment citées, voici quelques-uns des facteurs qui jouent un rôle important dans l'acquisition de la prononciation:

1. Attention, prise de conscience d'un aspect phonétique
2. Compréhension d'un principe
3. Nombre de fois entendu
4. Nombre de fois répété (et bien prononcé)
5. Côté visuel-auditif (voir la bouche et le visage du locuteur natif quand il prononce)

6. Côté visuel (voir un mot écrit, une liaison marquée en couleur)
7. Corrections (permettant le stockage et la fossilisation des bonnes formes)
8. Côté affectif (sentiments qu'ont les apprenants face à un cours, une matière, un instructeur)
9. Méthodes pédagogiques (analytiques)
10. Pratique des nouvelles formes (exercices structuraux, guidés ou semi-guidés)
11. Mise en pratique des nouvelles formes dans des activités de communication
12. Musique, chansons, lecture à haute voix, jeux...
13. Autres sens (toucher, odorat...) ; ou mélange de plusieurs sens et activités

Lors de la petite leçon de phonétique, nous nous sommes efforcés de prendre en compte autant de facteurs que possible dans l'explication théorique, en prenant en compte notamment les différents types d'intelligences que décrit Gardner, H. (1993). L'approche pédagogique choisie était celle de Long et Robinson (1998 :16) « le focus sur la forme », c'est-à-dire une approche plus traditionnelle, explicative, analytique. Nous avons expliqué pourquoi il est important d'enchaîner les mots, le principe de la syllabation ouverte et de la contrainte anti-hiatus du français. Pour nous assurer que tous les étudiants comprennent les explications théoriques, les explications ont été faites en anglais, étant donné que certains étudiants de première année ne comprennent pas tous le français s'il est parlé rapidement.

Pour l'explication plus détaillée, nous avons donné des exemples d'enchaînements en anglais et nous avons volontairement coupé les mots en anglais pour montrer l'effet produit si l'on n'enchaîne pas les mots. Nous avons donné de nombreux exemples de chaque catégorie en français, donnant aux étudiants une partie de l'input dont ils ont besoin, ce qui était particulièrement important pour les mots potentiellement inconnus présents dans le test de lecture, notamment pour les h aspirés. Pour les liaisons interdites, nous avons fait un signe du

doigt et un mouvement de la tête signifiant « non », indiquant que la liaison ne se faisait pas et nous avons aussi écrit au tableau ce qui était important.

Sur les photocopies (en annexe 6) distribuées pour la leçon de phonétique, une consonne entre parenthèses a été rajoutée pour toutes les liaisons (*un (N) avion, un (N) accent ; des (Z) enfants...*), aidant ainsi les apprenants visuels à se souvenir qu'il fallait prononcer la consonne de liaison avec le mot2.

Lorsque nous avons lu les photocopies distribuées, insisté sur les consonnes de liaison en les accentuant ; nous en avons exagéré l'intonation et l'intensité, tout en précisant que ceci était seulement pour l'explication, et que sinon, on les prononce « normalement », pour que les apprenants auditifs puissent se souvenir des points saillants.

A chaque fois qu'une liaison était faite, nous avons fait un geste de la main, de la gauche vers la droite, pour signaler que la consonne latente du mot1 s'enchaîne avec le mot2, ce qui pouvait aider tous les types d'apprenants, en particulier les apprenants kinesthésiques.

Nous avons écrit au tableau les phénomènes de liaison en écrivant quelques séquences avec un trait liant les deux mots, tout en disant les séquences à haute voix. Cela récapitulait tout ce qui avait été dit et pouvait profiter à tous les types d'apprenants.

Si les études précédentes reconnaissent l'importance de certains facteurs dans l'acquisition des apprenants, elles n'indiquent ni les facteurs qui sont plus importants que d'autres, ni ne précisent ceux qui améliorent le plus la production des apprenants. En effet, elles indiquent par exemple que l'instruction formelle utilisant telle ou telle méthode (des dictées, des logiciels, etc.) a produit de meilleurs résultats que ceux du groupe témoin qui ne l'a pas utilisé. Toutefois, aucune étude, à ma connaissance, ne compare les effets de certaines composantes de l'enseignement, pour déterminer par exemple si la répétition est plus utile que les corrections ou les explications.

Si l'instruction formelle joue un rôle déterminant, quelles activités sont les plus prometteuses : l'explication, la répétition ou la correction ? Ceci est une question cruciale car vu le nombre d'activités possibles, il importe de déterminer celles qui sont les plus efficaces et devraient alors être utilisées en priorité en classe.

Avec une approche analytique, non seulement les apprenants doivent être capables d'entendre et de produire certains « tronçons » (*chunks*), ils doivent aussi et surtout comprendre et internaliser un principe pour être capable d'appliquer ce qu'ils ont appris, avec d'autres mots ou groupes de mots qu'ils n'ont pas entendus pendant le cours.

Pour arriver à comprendre et internaliser un principe, les apprenants ont besoin au minimum de comprendre une règle, d'entendre et de répéter plusieurs exemples de mêmes formes (ex. un\_enfant, un\_arbre), d'être corrigés s'ils font des erreurs.

Pour tester le rôle des explications formelles, nous voulions déterminer les effets d'un facteur à la fois. Pour cela, nous avons quatre groupes de même niveau intermédiaire, tous inscrits dans le même cours. Tous les groupes ont reçu la même formation, avec le même professeur (moi-même), mais le test a eu lieu à des moments différents de leur formation, testant ainsi les effets des répétitions, des explications et des corrections. Ils ont reçu, pour cette étude, le même matériel (photocopies distribuées) et les mêmes exercices, employant la même approche « traditionnelle » : explication théorique, répétition, pratique et correction.

Dans les groupes 1, 2, 3, le test se faisait après certaines activités. Dans le groupe témoin, le groupe 4, le test s'est fait au début du cours, avant les activités, pour déterminer si les étudiants avaient fait des progrès simplement parce qu'ils avaient lu le même texte plusieurs fois et s'étaient familiarisés avec ses éléments, ou même d'autres facteurs, c'est-à-dire des facteurs autres que l'explication, la répétition ou la correction.

Les tests servent à déterminer quelle activité a joué le rôle le plus important dans l'acquisition des liaisons, et si elles ont toutes joué un rôle également important (ou n'ont pas

joué de rôle). L'ordre des activités de chaque groupe est résumé dans le tableau 3.4. Le moment de l'année et de la journée sont aussi précisés, parce qu'il est possible que le moment de la journée joue un rôle dans les progrès des étudiants.

Groupe 1 Le lundi 9 février 12h30-13h30	Groupe 2 Le lundi 9 février 13h30-14h30	Groupe 3 Le lundi 9 février 14h30-15h30	Groupe 4 Le lundi 9 février 15h30-16h30
	Très brève explication de ce qu'il faut faire 2mn	Très brève explication de ce qu'il faut faire 2mn	Très brève explication de ce qu'il faut faire 2mn
1. Explications 15 mn	1. Répétitions 15 mn	1. Répétitions 10 mn	1. TEST 10 mn
2. Répétitions 10 mn	2. TEST 10 mn	2. Corrections 10 mn	2. Explications 13 mn
3. Corrections 5 mn	3. Explications 8 mn	3. TEST 10 mn	3. Répétitions 10 mn
4. TEST 10 mn	4. Corrections 5 mn	4. Explications 8 mn	4. Corrections 5 mn

**Tableau 3.4. Horaire des activités de l'étude expérimentale**

Un chronomètre électronique (qui sonne doucement) assurait que chaque partie était bien enseignée dans les délais prévus. Toutefois, on peut remarquer que certains groupes avaient reçu plus plus de minutes d'enseignement que d'autres au moment du test 2 (le premier test a été administré la semaine précédente). Au final, tous les groupes ont reçu le même temps d'enseignement : 30 minutes.

### **La partie explication** (comprendre une règle)

Nous avons expliqué ce qui était sur les photocopies (voir annexe 6): les enchaînements puis les liaisons, en expliquant pourquoi on les a faits (règle anti-hiatus, préférence pour une segmentation du type CV), l'importance de le faire et nous avons donné les exemples de la feuille.

## **La partie répétition**

Nous avons lu les groupes de mots des photocopies et les étudiants ont répété tout de suite après en les lisant en même temps sur les photocopies distribuées. En écoutant les étudiants, nous avons remarqué les erreurs et cela a été mentionné dans la partie suivante (correction).

## **La partie correction**

Nous avons attiré l'attention des étudiants sur les erreurs commises, et expliqué que les séquences non liées, ou non-enchaînées ne sont pas une forme francophone du français standard. Nous avons donné la bonne forme et demandé aux étudiants de répéter deux fois chaque (bonne) forme corrigée qui avait causé quelques difficultés. Les principaux problèmes étaient les cas de non-resyllabation (liaisons non-enchaînées).

La formation et les tests devaient être courts pour ne pas trop changer le contenu et la forme du cours de laboratoire. Étant donné qu'une partie du cours était prévue pour les liaisons (le curriculum avait été décidé avant que nous ne contactions le directeur du cours), le contenu de la présentation était adéquat. Par contre, il ne fallait pas dépasser le temps qui était prévu à cet effet. Répondre aux questionnaires, faire les enregistrements (pré-test, test et post-test) a déjà demandé du temps aux étudiants volontaires. Il était impossible de leur donner une formation plus longue.

Les explications, les répétitions et les corrections ont eu lieu dans un temps très limité pour enseigner un sujet aussi vaste que les liaisons et les enchaînements. Idéalement, il aurait fallu au moins plusieurs heures avec des exercices beaucoup plus variés, notamment des exercices de perception, des dictées, des exercices guidés, semi-guidés et des exercices qui mettraient en pratique les éléments étudiés dans des activités de communication. Nous ne



pouvions pourtant pas bouleverser l'emploi du temps du cours de laboratoire et avons dû nous restreindre à ce temps limité.

C'est la raison pour laquelle seules les liaisons obligatoires et interdites ont été enseignées, les enchaînements et liaisons facultatives ont seulement été mentionnés dans l'explication sur la contrainte anti-hiatus et le système de syllabation ouverte.

Toutefois, cette restriction de temps était aussi l'un des défis prometteurs de cette étude. Si les étudiants pouvaient faire des progrès après seulement 30 mn de cours, on peut facilement imaginer les progrès qu'ils feraient après une année entière de cours de phonétique corrective ou au moins 40-50 mn chaque semaine dans un cours de langue d'un semestre ou deux.

Les buts de cette étude sont de déterminer d'une part la performance des étudiants en première année universitaire de niveau intermédiaire ayant suivi au moins six années de français quant à la réalisation des liaisons, d'autre part de déterminer les progrès qu'ils ont faits, et enfin d'analyser les facteurs qui ont contribué le plus à leurs progrès, c'est-à-dire les effets de certaines techniques pédagogiques. Pour limiter les effets d'autres facteurs, le contenu et matériel présenté (nouveau pour tous), le professeur et la méthode pédagogique (traditionnelle) étaient les mêmes dans tous les cours.

Toutefois, les facteurs propres à l'apprenant lui-même sont des facteurs indépendants de l'enseignement et de l'enseignant. Ces facteurs sont, d'après Suter (1976) et Purcell et Suter (1980), les facteurs qui prédisent le plus le succès des apprenants. D'après Purcell et Suter (1980), les quatre facteurs prédisant le mieux le succès des apprenants sont la L1, les capacités d'imitation, et plusieurs variables réunies en rapport avec le temps passé dans le pays de la langue cible et la volonté (le désir ou le besoin) d'améliorer la prononciation.

Afin d'évaluer les facteurs personnels dont parlent Purcell et Suter (1980) qui participent à leurs progrès, nous avons demandé aux participants de remplir un questionnaire

sur leur formation linguistique, leurs impressions sur leurs connaissances, leur expérience linguistique et d'autres facteurs personnels pour déterminer s'il y avait une corrélation entre ces facteurs et leurs progrès. Pour plus de détail, voir le questionnaire en annexe 3.

## **6. Le texte à lire**

Puisque le texte à lire constitue la base de notre recherche, il était primordial que le texte cible les points que nous souhaitions étudier et réponde à des exigences précises. Les textes existant déjà qui ont été créés pour des francophones et qui n'ont pas été créés pour une étude phonétique particulière risquent de ne pas avoir assez d'items pour évaluer une catégorie. Par exemple, le texte 1 de Mastromonaco ne contient que 6 contextes [dét. + nom], et 5 de ces contextes sont probablement acquis depuis longtemps chez des sujets ayant étudié le français au moins 6 ans. Les textes qui n'ont pas été créés spécialement pour des études sur les liaisons vont forcément en comprendre un plus petit nombre qu'un texte « artificiel » spécialement créé à cet effet. Au risque que le texte ne soit pas très authentique, la meilleure option nous semble être de créer un texte contenant plusieurs éléments à tester.

Étant donné le nombre considérable de contextes de liaisons obligatoires avec des sous-catégories pour chaque catégorie, nous nous sommes concentrés en priorité sur les liaisons obligatoires dans le groupe nominal. Les liaisons facultatives, interdites et les enchaînements, en nombre plus restreint sont présents dans le texte afin de confirmer (ou d'infirmer) les hypothèses concernant les liaisons obligatoires.

Notre étude a pour but de tester les compétences des apprenants anglophones du français L2. Pour cela, ils ont lu un texte, et plusieurs groupes de francophones ont lu le même texte. Nous avons ainsi pu comparer les productions des francophones et des anglophones.

Nous avons choisi d'utiliser les classifications traditionnelles de Delattre (1951) qu'utilisent la plupart des chercheurs dans ce domaine (Encrevé, Mastro Monaco, Boula de Mareüil, les chercheurs qui participent au projet PFC, etc.) ainsi que l'Académie Française, afin de déterminer des catégories globales de liaisons qui devraient être réalisées (obligatoires), ne devraient pas être réalisées (interdites) et d'évaluer aussi les productions de liaisons facultatives et d'enchaînements.

Nous souhaitons étudier principalement les facteurs linguistiques présents dans le texte à lire dans la production de liaisons. Pour cela, nous avons considéré les facteurs syntaxiques, phonétiques et lexicaux du texte que les sujets devaient lire. Nous nous sommes concentrés dans cette thèse sur le rôle de la fréquence dans le cadre du MBU, sur les productions de liaisons et enchaînements des anglophones de niveau intermédiaire (A2-B1), comparées aux productions de francophones lisant le même texte.

Le texte est séparé en phrases pour en faciliter la lecture. Nous nous sommes efforcés d'éviter les mots difficiles, peu fréquents, ou longs. Le texte a été créé de façon à réunir les phénomènes de sandhi les plus fréquents et regroupe un total de 13 enchaînements, 51 liaisons obligatoires, 17 liaisons facultatives et 14 liaisons interdites soit un total de 95 cas de sandhi possibles sur 236 mots du texte. Le texte contient une situation « liante » tous les 2.48 mots, un enchaînement tous les 18 mots, une liaison obligatoire tous les 4 mots, une liaison facultative tous les 13 mots et une liaison interdite tous les 16 mots.

A titre indicatif, le texte à lire dans le projet PFC contient un total de 35 sites de liaisons. Toutefois le projet PFC prévoit également deux situations de parole spontanée. Puisque pour notre étude nous avons seulement un texte à lire et pas de parole spontanée, il nous paraissait nécessaire de réunir une plus grande variété de contextes.

Le texte a été conçu pour tester certains contextes que nous pensons être parmi les plus facilement acquis par les étudiants : les liaisons dans le groupe nominal, et en particulier

(mais pas seulement) dans le contexte [dét. + nom]. En effet, c'est ce contexte syntaxique que les enfants francophones apprennent en premier (Dugua, 2005, 2006) parce que les articles et les noms sont plus fréquents que les autres mots (prépositions, conjonctions, adverbes).

D'une part, les contextes choisis sont les plus fréquents : [art. + nom], [adj. + nom], etc. D'autre part, nous supposons que les séquences choisies ont été très souvent entendues et prononcées par les apprenants. Dans notre texte, d'après *Vocabprofil*, 87.70% (214 mots) des mots sont de fréquence K1, les 1000 mots les plus fréquemment utilisés -1 à 1000, (*amis, enfants, histoires, intelligent, arbre...*) ; 5.74% (14 mots) appartiennent à la catégorie K2, les 1000 mots suivants les plus utilisés (1001 à 2000), 1.23% (3 mots) à la catégorie K3, les 1000 mots suivants les plus utilisés (2001 à 3000), et 5.33% (13 mots) à la catégorie Hors-Liste (*Off-list words*) abrégé en HL : les mots qui ne sont présents dans aucune des trois catégories précédentes, c'est-à-dire à partir de 3001.

Parmi les mots appartenant à la catégorie HL (*acclamations, affreux, aigu, ennui, éponges, gigolos, habits, handicaps, hollandais, hongrois, oies, yachts, yogourts*), beaucoup existent en anglais, ou sont proches de l'anglais. En d'autres termes, les mots plus rares ne sont pas complètement inconnus des anglophones : *acclamations, gigolos, handicaps, hollandais, yachts, yogourts*. Les mots *éponges* et *Hongrois* sont proches de *sponge* et *Hungarian*. Tout a été fait pour tester des contextes différents (h muets, h aspirés, semi-consonnes, nasales, etc.), mais de façon à ce que la plupart des mots soient le plus susceptibles d'être connus des étudiants.

Cette étude comportait plusieurs défis, notamment en raison des restrictions de temps. Pour arriver en un temps réduit à créer un enregistrement représentatif, il fallait élaborer un texte assez court, mais aussi un texte riche et « concentré » en liaisons et enchaînements, qui puisse tester au mieux différents contextes phonétiques, dans des phrases qui aient du sens avec des mots connus des étudiants.

Il y avait un dilemme. Il fallait à la fois éviter les mots rares et inconnus qui pouvaient créer des problèmes particuliers aux étudiants, mais à la fois tester différents contextes tels que les h aspirés et muets, les semi-consonnes etc., qui se trouvent toujours dans des mots moins fréquents. Il fallait aussi que certains mots soient plus rares que d'autres afin de tester les effets de fréquence et que des mots soient plus longs afin de tester l'effet de longueur des mots sur la production des étudiants.

Le but n'était pas de piéger les étudiants avec des contextes particulièrement difficiles pour eux, mais de tester certains éléments, en séparant, autant que possible, les difficultés. Nous avons donc séparé les difficultés : longueur des mots, fréquence des mots, familiarité des mots. Des mots peuvent être rares (peu fréquents dans l'échelle de fréquence de *Vocabprofil*), mais tout de même familiers des étudiants s'ils sont proches de l'anglais. C'est dans cet effort par exemple que le mot *yogourt* a été choisi (plutôt que *yaourt*), mot plus fréquemment utilisé chez les francophones canadiens, et plus proche de l'anglais (*yogurt*), mot que les étudiants auraient pu trouver sur les pots de yaourts.

Les difficultés de lecture en ce qui concerne les mots eux-mêmes étaient de lire soit des mots potentiellement inconnus mais assez courts (*ennui, éponge, hongrois*), soit des mots longs mais connus (*accident, adapter, handicaps, aventures, intelligent, édifice, hollandais* etc.). En principe, les étudiants de ce niveau intermédiaire devraient connaître la plupart des mots de trois ou quatre syllabes du texte: *demandé, invités, étudiants, assistants, s'adapter, édifice, attendu, acclamations, aventures, gigolos, handicaps, hollandais, intelligent, accident*.

En outre, les mots potentiellement difficiles à prononcer ou potentiellement inconnus ou moins fréquents étaient présents dans la présentation formelle, pour que s'ils étaient inconnus ou rarement entendus, ils ne soient plus totalement inconnus après la présentation, telles que les expressions figées : *fait accompli, accent aigu*.

Cette étude se concentrant surtout sur les liaisons obligatoires dans le groupe nominal dans le contexte [dét. + nom] et [nom + adj.], nous avons choisi des expressions figées qui rentraient dans ce contexte : *accent aigu*, *sous-entendu*, etc.

Les étudiants de ce niveau devraient tous avoir entendu de très nombreuses fois les expressions *accent aigu* et *États-Unis*, plusieurs fois *sous-entendu* et un peu moins souvent *devant le fait accompli*. Toutefois, les expressions figées dans le groupe nominal se trouvant en nombre limité, nous avons quand même utilisé *le fait accompli*. Cette expression a été présentée et répétée plusieurs fois en classe dans la petite présentation.

Les liaisons facultatives ont été choisies pour le vaste éventail qu'elles représentent. Certains contextes [aux. + part. passé] et [nég. + part.passé] sont considérés comme étant des liaisons facultatives fréquentes d'après Malécot (1975), alors que d'autres sont plus rares, notamment avec le r de l'infinitif (*s'adapter à*).

D'après la théorie de la fréquence, les liaisons acquises en priorité sont celles qui sont le plus fréquemment entendues. Nous nous attendons à ce que les étudiants ne fassent que les liaisons obligatoires et quelques liaisons facultatives les plus entendues.

En ce qui concerne les enchaînements, il s'agissait ici de tester les contextes les plus fréquents : [pron. + verbe], [dét. + nom], [nom + verbe] et [chiffre + nom]. Ils sont fréquents à l'oral, en particulier les pronoms *il* et *elle* + voyelle.

La raison pour laquelle l'article *une* n'a pas été choisi, alors qu'il est pourtant le cas le plus fréquent d'enchaînement, est parce que si l'article est mal prononcé (/œ̃/ou /œ̃n/) ou si le *n* n'est pas complètement prononcé, cela pourrait empêcher l'enchaînement ou devenir une liaison. Alors qu'avec *cet*, il n'y a généralement aucune difficulté de prononciation du /t/. De plus, cet adjectif démonstratif est aussi très fréquent. Là aussi, d'après la fréquence de ces éléments, on peut en déduire que, si l'apprenant ne fait pas l'enchaînement après *il*, *elle*, *cet* et *cette*, il ne le fera probablement pas ailleurs.

Les francophones et les anglophones ont lu un texte de 236 mots. Les consignes étaient : « Veuillez lire ce texte à vitesse normale ».

### **Les contextes phonétiques et syntaxiques**

En ce qui concerne les contextes syntaxiques (voir annexe 8), il y a dans le texte différentes sortes d'articles (défini, indéfini, singulier, pluriel) et d'adjectifs (numéraux, cardinaux, possessifs, qualificatifs), à différentes personnes et genre (masculin singulier, masculin pluriel, féminins pluriels). Le texte comprend différents contextes phonétiques: liaisons avec nasales, consonnes latentes du masculin similaires au féminin, semi-voyelles, h aspirés, h muets, etc.

Enfin, les mots sont de fréquence différente, bien que la plupart soient des mots très fréquents. Nous nous sommes efforcés de trouver des contextes phonétiques et syntaxiques légèrement différents, bien qu'appartenant en majorité au groupe nominal, sauf le cas de « dans un » (préposition + article) pour les liaisons obligatoires.

Nous estimons que 95 cas de sandhi, dont 81 dans le groupe nominal, devraient permettre d'évaluer les compétences des apprenants quant à la liaison et à l'enchaînement dans le groupe nominal.

## **7. Fréquence lexicale et banque de données**

L'un des facteurs de fréquence que nous voulons tester est la fréquence des mots et le rôle de cette fréquence sur la production de liaisons. Pour cela, nous avons établi un indice de fréquence des mots<sup>1</sup>, des mots<sup>2</sup> et de cooccurrence de ces mots avec la banque de données de *Lexique 3*, puis avec la banque de données de *Vocabprofil*. Examinons pourquoi nous avons choisi d'utiliser ces banques de données, de rechercher certaines informations et d'établir ces

indices et comment nous les avons attribués et expliquons ce que représentent ces bases de données.

Dans une séquence de deux mots, le mot2 joue un rôle important dans la production de liaisons, puisqu'il peut déterminer si une liaison se fait ou non dans un même contexte syntaxique. Par exemple, on fait la liaison avec *les\_enfants*, mais pas avec *les / hanches* ou *les / yaourts* alors que le mot1 est le même (*les*). De plus, le mot2 est souvent plus représentatif de la fréquence lexicale que le mot1. En effet, bon nombre de mots1 sont des déterminants et le nombre de mots1 possibles est souvent plus restreint que le nombre de mots2 possibles. En revanche, le mot1 est fortement lié à la fréquence syntaxique, car c'est lui qui détermine en partie le type de structure syntaxique. Par exemple, si le mot1 est *les*, cela indique que la structure syntaxique est [dét. + nom (ou adj.)].

En outre, le mot1 n'est pas seulement représentatif d'une structure syntaxique, il fournit aussi d'autres informations. Tout d'abord, il donne des informations sur des sous-parties des structures syntaxiques. Par exemple, à l'intérieur de la structure [dét. + nom], on compte article défini, article indéfini, adjectif possessif, adjectif démonstratif, etc.

De plus, le mot1 joue un rôle tout aussi important que le mot2 puisque, d'une part la liaison dépend de la consonne de liaison, d'autre part, la liaison dépend directement du mot1 et de la fréquence de celui-ci, et enfin, le mot1 fournit d'autres informations telles que le genre, le nombre, la personne, etc. Par exemple, on fait la liaison avec *des\_histoires*, mais beaucoup moins avec *premières\_histoires* alors que le mot2 est le même (*histoires*).

Il apparaît donc que trois facteurs rentrent en compte quant à la fréquence des mots formant la liaison : la fréquence du mot1, la fréquence du mot2 et la fréquence de cooccurrence du mot1 et du mot2.

La fréquence de cooccurrence diffère de la fréquence du mot1 et de la fréquence du mot2, car c'est elle qui détermine la « cohésion » de la fréquence de la séquence. Au niveau



du traitement des données, ce type de fréquence joue un rôle majeur. En effet, les locuteurs habitués à entendre certains mots suivis d'autres mots seront plus enclins à les traiter comme une seule unité si leur taux de cooccurrence est élevé. Cela devrait avoir pour conséquence de faciliter la production de la liaison.

Nous pensons que la fréquence d'utilisation du mot1, la fréquence du mot2 et la fréquence de cooccurrence des mots sont interreliés et jouent ensemble un rôle déterminant pour l'acquisition et la réalisation des liaisons et enchaînements. En effet, si deux mots sont très fréquents pris isolément, il y a plus de chances qu'ils apparaissent ensemble (que leur taux de cooccurrence soit élevé), bien que ce ne soit pas systématiquement le cas. Si un mot est entendu et utilisé très fréquemment, il est bien ancré dans le lexique du locuteur et il y a plus de chances pour que la liaison soit produite. Pourtant, même si deux mots sont relativement rares comme par exemple, *dangereux individu*, si leur taux de cooccurrence est assez élevé, la liaison aura plus de chances d'être produite que si deux items sont peu fréquents et que leur taux de cooccurrence est très faible ou nul.

Mais pour évaluer la fréquence des mots, c'est-à-dire la fréquence à laquelle les mots sont entendus et prononcés, il importe d'utiliser une banque de données portant sur la langue orale. Or, trouver un corpus de langue orale représentatif de la parole spontanée, n'est pas chose facile. Comme l'indiquent Cappeau et Gadet (2007), il est difficile de trouver des corpus de français parlé pour plusieurs raisons. La plupart des corpus sont ceux récoltés par des personnes isolées pour leurs recherches (doctorat). Seul un cercle restreint de personnes connaissent l'existence de ces corpus et comment y accéder. Les corpus sont le plus souvent créés pour répondre aux questions qui intéressent directement les chercheurs, ce qui rend les corpus très spécialisés et souvent moins utiles aux chercheurs qui s'intéressent à d'autres phénomènes.

Un autre problème est celui de l'accessibilité. Si les ressources sont accessibles mais difficiles d'utilisation, elles sont en définitive peu accessibles. Si des recherches sur l'internet ne permettent pas d'accéder à ces corpus et qu'il faut avoir l'adresse du site web au préalable pour y accéder, il faut connaître les auteurs et leurs recherches pour pouvoir trouver le site. Ainsi les corpus et recherches correspondants doivent être connus de la communauté de chercheurs, pour que derniers sachent qu'ils existent, en connaissent le nom et puissent ainsi faire des recherches.

En outre, les plus grands corpus, et les corpus les plus nombreux sont ceux de l'anglais. En France, il y a surtout eu des corpus écrits tels que celui de *Frantext*.

Enfin, ce n'est que récemment qu'ont apparu des projets accessibles à tous, utilisables sur l'internet, qui sont assez riches en informations pour intéresser des chercheurs de divers horizons, ce qui est le cas par exemple du corpus PFC. Comme le mentionnent Cappeau et Gadet (2007 : 130) : « [...] Le corpus PFC comporte 30 points d'enquête répartis dans l'ensemble de la France. Chaque locuteur a été enregistré en lecture et en conversation libre. C'est certainement le plus gros corpus oral de français, étiqueté et consultable (1.400.000 mots, transcrits pour moitié). »

New et al. (2007) expliquent également les difficultés à réunir un corpus de français oral. Étant donné qu'il est difficile de mettre sous forme écrite la parole orale, car c'est très coûteux en termes de temps et d'argent, une solution était d'utiliser les sous-titres de films français qui correspondent aux paroles des acteurs. Vu que ce sont le plus souvent des dialogues, la parole est censée être authentique et représentative de la langue orale de tous les jours.

De plus, c'est celle que les locuteurs entendent régulièrement à la télévision ou au cinéma. C'est ainsi que la banque de données de *Lexique* (du CNRS) a créé une banque de données de la langue orale. L'équipe a obtenu les sous-titres de 9474 films et séries de

télévisions pour lesquels il y a quatre catégories : les films en français, les films en anglais, les séries télévisées en anglais et les films d'autres langues européennes.

Pour notre étude, nous avons besoin d'une banque de données qui nous permette d'évaluer la fréquence de certains mots à l'oral et la fréquence de leur cooccurrence. Le site de *Lexique 3* est facilement accessible et fournit les renseignements dont nous avons besoin. La banque de données *Lexique 3* se base sur les recherches de New, Pallier, Ferrand et Matos (2001)<sup>40</sup>. Nous avons utilisé la banque de données de *Lexique 3* qui contient les films sous-titrés en français. Cela constitue un total de 1.9 million de mots (basés sur des films tels que *Camille Claudel*, et *C'est arrivé près de chez vous*).

Comme les auteurs le mentionnent eux-mêmes, les sous-titres comportent quelques restrictions : les phrases sont souvent raccourcies (pour que les locuteurs puissent lire rapidement les sous-titres) et certains sujets sont parfois surreprésentés (sujets « populaires » tels que la police, les crimes, les relations amoureuses, etc.), au détriment d'autres thèmes pourtant présents dans la vie réelle. Toutefois les auteurs ont comparé les résultats obtenus avec la banque de données de *Lexique 3* avec d'autres banques de données et ont trouvé que les résultats étaient fiables (New et al. 2007).

Pour notre étude, nous voulions utiliser des banques de données les plus représentatives possibles de l'expérience des locuteurs francophones dans ce qu'ils ont pu entendre et lire. Pour cela nous avons opté pour deux solutions : utiliser les banques de données de *Lexique 3* et de *Vocabprofil*. Celles-ci fournissent des informations différentes et sont basées sur des corpus différents. Nous pensons ainsi obtenir des renseignements représentatifs de la fréquence de certaines séquences.

Lire à haute voix met en œuvre plusieurs compétences. Le locuteur doit être capable de reconnaître phonétiquement les mots et séquences. Il doit avoir appris comment les mots et

---

<sup>40</sup> Le site web d'où les données sont accessibles est : <http://www.lexique.org>

séquences se prononcent, les avoir ancrés pour être capables de les prononcer correctement. D'autre part, le locuteur doit être capable de reconnaître les séquences orthographiquement. Il doit pouvoir reconnaître les mots écrits, en comprendre le sens et aussi être capable d'évaluer inconsciemment et rapidement les groupes syntaxiques, prosodiques et phonétiques pour pouvoir les lire sans trop d'hésitations. Cela demande une certaine expérience en lecture, donc des connaissances du français écrit.

Puisqu'il faut connaître les formes orales pour savoir les prononcer, nous avons choisi d'utiliser *Lexique 3* (portant sur les sous-titres des films français). Mais puisqu'il faut aussi reconnaître les mots de l'écrit, il importe de prendre en considération la fréquence des mots à l'écrit, c'est-à-dire une banque de données de fréquence de l'écrit. En effet, les sujets et les mots de l'écrit peuvent différer de ceux de l'oral. Par exemple, en interrogeant *Lexique 3* on trouve que *est* a un taux de fréquence de 19 417 (par million de mots) dans les films et de 6331 dans les livres.

La négation *pas* a un taux de fréquence de 18 188 (par million de mots) dans les films et de 8795 dans les livres. Rappelons qu'il est fréquent d'omettre la négation *ne* à l'oral, mais impossible d'omettre la négation *pas*. Par contre, à l'écrit, on peut omettre la négation *pas* avec certains verbes (ex : *on ne peut imaginer*), bien que soit assez rare, alors qu'on ne peut en aucun cas omettre la négation *ne* à l'écrit. Les normes de l'oral et de l'écrit sont différentes, d'où l'intérêt de trouver des banques de données qui les examinent séparément.

D'autre part, la façon dont la fréquence est calculée dans *Lexique 3* et dans *Vocabprofil* est différente. Thomas (2008) explique l'origine et l'utilisation de *Vocabprofil* :

[...] Le logiciel *Vocabprofile* de Laufer et Nation (1995) a permis de mesurer la richesse lexicale de textes écrits en anglais, en calculant la proportion des mots les plus rares, c'est-à-dire ceux qui dépassent le 2000e rang dans les listes de fréquence. Il a été adapté pour le français par Jones (2001), puis validé et utilisé d'abord par Goodfellow et Lamy (2002) sur un corpus de français L2 faible, puis par Ovtcharov et al. (2006) sur un corpus de français L2 avancé, fourni par des fonctionnaires Anglophones canadiens. Ces deux études concluent qu'on peut évaluer la richesse lexicale des textes français dès le 1000e rang de fréquence. [...] (Thomas, 2008 : 588)

*Vocabprofil* calcule la fréquence d'utilisation des mots d'après les critères suivants :

K1 : les 1000 mots les plus fréquemment utilisés (1 à 1000)

K2 : les 1000 mots suivants les plus utilisés (1001 à 2000)

K3 : les 1000 mots suivants les plus utilisés (2001 à 3000)

Mots Hors-liste (*Off-list words*) abrégé en HL : les mots qui ne sont présents dans aucune des trois catégories précédentes, c'est-à-dire à partir de 3001.

Alors que dans *Lexique 3* la fréquence des items est calculée à partir d'un million de mots, dans *Vocabprofil* la fréquence est calculée selon la fréquence relative à d'autres mots. Par exemple, le mot *elle* dans *Lexique 3* a un taux de fréquence de 4520, auquel nous avons attribué le code j, et *il* a un taux de fréquence de 13 222, auquel nous avons attribué le code m.

Par contre, dans *Vocabprofil*, le mot *elle* appartient à la catégorie des 1000 mots les plus fréquents, c'est pourquoi nous lui avons attribué l'indice 3, et *il* appartient également à la catégorie des 1000 mots les plus fréquents, c'est pourquoi nous lui avons aussi attribué l'indice 3. Autrement dit, d'après *Vocabprofil*, leur taux de fréquence est le même, alors que dans *Lexique 3*, on remarque une légère différence de fréquence entre *il* et *elle*.

De même, le mot *accident* (présent dans notre texte) apparaît 100 fois dans *Lexique 3*, et nous lui avons attribué le code f, alors qu'il appartient pourtant à la catégorie des 1000 mots les plus fréquents dans *Vocabprofil*, c'est pourquoi nous lui avons attribué l'indice 3.

Un mot peut appartenir à la catégorie des 1000 mots les plus fréquents dans *Vocabprofil*, mais cela ne veut pas dire qu'il soit pour autant employé le plus grand nombre de fois dans *Lexique 3*. C'est la raison pour laquelle nous avons choisi d'utiliser la banque de données de *Lexique 3* pour évaluer la fréquence des mots de l'oral, ET aussi la banque de données de *Vocabprofil* pour évaluer la fréquence des mots à l'écrit. Nous pouvons de cette façon évaluer et considérer deux types de fréquences.

Ainsi, nous avons analysé le rôle de différents facteurs de fréquence : la fréquence lexicale, la fréquence des structures syntaxiques (*type frequency*), la fréquence des consonnes de liaison (/n/, /t/, /z/, /R/), et les avons comparés aux résultats obtenus avec l'étude du corpus PFC, dans l'étude de Mallet (2008) et de Durand et Lyche (2008).

Nous avons analysé la fréquence des mots de notre étude (de notre texte à lire) avec la banque de données de fréquence de *Lexique 3* (qui est un corpus de langue orale), et celle de *Vocabprofil* (qui est un corpus de langue écrite). Ces deux bases de données nous permettent de déterminer la fréquence lexicale suivante :

1. la fréquence du mot1 avec *Lexique 3*
2. la fréquence du mot2 avec *Lexique 3*
3. la fréquence de cooccurrence des mot1 et mot2 avec *Lexique 3*
4. la fréquence du mot1 avec *Vocabprofil*
5. la fréquence du mot2 avec *Vocabprofil*

Nous étudierons un peu plus loin le rôle des différents types de fréquence dans la production de liaisons, ainsi que la productivité et la complexité des structures. Nous analyserons le rôle de la fréquence dans les productions des francophones et des anglophones.

### **7.1.1. La banque de données de *Lexique 3***

Afin de mieux comprendre l'effet de la fréquence sur la production de liaisons, nous avons cherché la fréquence des mots isolés (fréquence du mot1 et fréquence du mot2) de chaque séquence potentiellement liante du texte de notre étude, puis de cooccurrence de ces mots. Pour cela, nous sommes allés à *Recherche par mot* puis à *Lexique 3.55*. Par exemple pour l'enchaînement *1cette Anglaise*, le résultat est le suivant :

mot1 (*cette*) = 1902.27

mot2 (*Anglaise*) = 51.01

Nous avons arrondi ces chiffres (ici 1902 et 51) pour tous les mots du texte. Ceux-ci correspondent à la fréquence par million de mots. Puis nous leur avons attribué un code, des lettres allant de *a* à *m* comme indice de fréquence pour les mots isolés :

### Fréquence d'occurrence des mots de *Lexique 3*

0 à 10	a	+ les mots sont rares
11 à 20	b	
21 à 40	c	
41 à 60	d	
61 à 80	e	
81 à 100	f	
101 à 200	g	
201 à 500	h	
501 à 1000	i	
1001 à 5000	j	
5001 à 6500	k	
6501 à 8000	l	
8001 et +	m	+ les mots sont fréquents

Par exemple, pour la séquence *cette Anglaise*, les résultats sont les suivants :

code	enchaînement	fréq mot1	ind
E1	cette	1902	j

**Tableau 3.5.** Fréquence mot1 *Lexique 3*

code	enchaînement	fréq mot2	ind
E1	Anglaise	51	d

**Tableau 3.6.** Fréquence mot2 *Lexique 3*

La banque de données calculant la cooccurrence des mots réunit des données différentes de celle des mots isolés. C'est la raison pour laquelle un mot a été trouvé en cherchant son taux de cooccurrence avec un autre mot, alors qu'il était introuvable lorsque cherché seul sur *Lexique 3*. Comme la fréquence de cooccurrence des mots est différente de celle des mots isolés, nous avons attribué les mêmes indices de fréquence (de *A* à *M*), mais pour des chiffres différents :

### Fréquence de cooccurrence des mots de Lexique 3

0 à 9	A	+ les mots sont rares
10 à 19	B	
20 à 39	C	
40 à 59	D	
60 à 79	E	
80 à 99	F	
100 à 199	G	
200 à 299	H	
300 à 449	I	
450 à 499	J	
500 à 649	K	
650 à 799	L	
+ de 800	M	+ les mots sont fréquents

Par exemple, le taux de cooccurrence de *cette Anglaise* est le suivant :

code	enchaînements	cooc.	cooc
E1	cette Anglaise	4	A

**Tableau 3.7. Fréquence cooccurrence *Lexique 3***

Nous avons choisi ce type de classification, contenant des lettres minuscules, des lettres majuscules (et autres codes pour les autres facteurs) pour les indices de fréquence parce que, pour notre codage dans *GoldVarb*, il fallait des codes différents (pour les mots isolés ou le taux de cooccurrence) afin d'éviter les erreurs possibles.

D'après le MBU, passé un certain nombre d'occurrences, la fréquence n'a plus d'importance. C'est-à-dire qu'un sujet ait entendu une séquence 1000 fois ou 5000 fois ne semble plus constituer de différences, vu qu'à partir d'un certain nombre d'occurrences entendues, la structure ou la séquence a été acquise. Par contre, ce sont les étapes intermédiaires qui semblent être plus représentatives de l'acquisition de structures et d'items. Aussi, découper les différentes fréquences par tranches permet un peu plus de précision qu'en découpant en plus grands groupes.



## Un mot manquant

Nous avons trouvé la fréquence de tous les mots du texte, sauf pour le mot *yogourts*. En faisant des recherches pour les mots les plus proches et ayant trouvé 4.73 pour *yaourts*, 0.31 pour *yogourt* (singulier), nous en avons déduit que si les mots les plus proches sont déjà très rares et que le mot recherché n'est pas disponible, c'est parce qu'il est encore plus rare. Nous pouvons donc lui attribuer le code *a* correspondant aux mots très rares.

### 7.1.2. Analyse des tableaux de *Lexique 3*

Si on analyse les résultats obtenus avec la fréquence des mots et de leur cooccurrence avec *Lexique 3* (voir annexe 9), on remarque plusieurs faits.

Tout d'abord, il existe de grandes disparités à l'intérieur d'une même catégorie grammaticale. Par exemple, parmi les pronoms personnels *il* a une fréquence maximale (13 222), alors que *elle* est trois fois moins fréquent (4520). Il en est de même avec les adjectifs possessifs qui, selon la personne, ont une fréquence très différente : *mon* (3848), *son* (1740), *mes* (1102), *tes* (690), *nos* (524), *leurs* (243).

Les tableaux en annexes 9, 10 et 11 indiquent que, comme on pourrait s'y attendre, les mots les plus fréquents sont globalement les verbes et auxiliaires *être* et *avoir*, l'adverbe *pas*, les pronoms (*il*, *elle*), les déterminants (*un*, *les*, *des*), les prépositions (*à*, *en*, *dans*), les adjectifs possessifs et démonstratifs, les mots généraux (*homme*, *personne*) et les adjectifs numériques et qualificatifs. Les mots les plus courts sont les plus fréquents en accord avec les résultats du PFC et les études de Durand et Lyche (2008) et Mallet (2008).

Sur les 51 mots les plus fréquents, seuls 6 mots (*aura*, *allés*, *attendu*, *enfants*, *enfant*, *entendu*) ont deux syllabes ou plus. Longueur de mot et fréquence sont des facteurs intimement liés. Comme le remarque Ågren (1973 : 30) : « [...] le facteur *longueur des mots*

*à lier* est en rapport avec le facteur *fréquence d'emploi* et que l'un et l'autre sont certainement capables d'influencer la liaison. »

Parmi les faits attendus, les auxiliaires sont plus fréquents que les verbes, puisqu'ils peuvent s'utiliser avec plus de verbes quand ils sont auxiliaires. Il ne semble pas y avoir ici de différence entre *a* et *ont* auxiliaire (18 559), ni entre le futur *aura* (13 572) et le présent. Par contre, on remarque une différence avec leur taux de cooccurrence. N'oublions pas que les banques de données sont différentes pour les mots isolés et les mots en couple. *Il a* a un taux de cooccurrence plus élevé (+ de 1000) que *il aura* (951) ; puisqu'il s'agit du même pronom *il*, cela signifie en fait que *a* est plus fréquent que *aura* (même si dans notre codage, nous les avons mis dans la même catégorie M, + de 800).

*Cet* est moins fréquent (499) que *cette* (1902), puisqu'il est « divisé » en deux selon que le mot suivant commence par une consonne (*ce*) ou une voyelle (*cet*), alors que le féminin reste le même indépendamment du phonème suivant.

Plus les mots sont fréquents pris séparément, plus on les trouve souvent ensemble (*il + a*, *homme + a*). De même, plus les mots sont rares, plus ils sont rarement en cooccurrence (*anglaise*, *accident*). Autrement dit, on peut prévoir que, si deux mots fréquents sont ensemble, leur taux de cooccurrence sera plus élevé que si deux mots moyennement fréquents ou rares sont ensemble. C'est ce qu'on observe avec les expressions figées. Par exemple, *plein air* se compose de deux mots moyennement fréquents : *plein* (208) *air* (485) et le taux de cooccurrence de ces deux mots est de 106. *Accent aigu* comprend deux mots rares : *accent* (14) et *aigu* (5). Leur taux de cooccurrence est de 1, c'est-à-dire qu'il est très rare.

Par contre, il y a aussi des mots rares qui assemblés dans une expression figée (qui au final ne constituent qu'un seul mot grâce au trait d'union), ont un taux de cooccurrence très élevé, comme c'est le cas avec *États-Unis* (+ de 1000).

Certaines expressions figées sont moyennement fréquentes : *sous-entendu* (38), *moyen-âge* (21) et *fait accompli* (12) et d'autres sont rares : *accent aigu* (1) et même très rares : *faux-airs* (0).

D'après le MBU, puisque toutes ces expressions figées n'ont pas la même fréquence, elles ne devraient pas être traitées de la même façon, ce dont nous reparlerons plus tard.

### **Les liaisons obligatoires**

Il est intéressant de remarquer que *un* est le plus fréquent (12 087) de tous les déterminants. Il est deux fois plus fréquent que son homologue pluriel *des* (6055) et aussi beaucoup plus fréquent que *les* (8720) et que *une* (7907). On peut attribuer cela au fait que *un* a deux sens : il peut être un article indéfini, mais aussi un numéral. De plus, l'initiale du mot ne change pas sa forme (contrairement par exemple à *le* qui devient *l'*). D'après le MBU, puisque les articles sont les plus fréquents et que *un* est le plus fréquent des articles, cela signifie que sa liaison devrait être la plus effectuée. Cela explique en partie le fait que la liaison en /n/ soit réalisée à 96,6 %.

Les adjectifs numéraux ne sont pas tous employés à la même fréquence : *deux* (1009), *trois* (380), *six* (117), *neuf* (59), *cent* (43). Mis à part *vingt* (29) et *dix* (118) qui peuvent être employés dans plusieurs nombres (*vingt et un*, *vingt-deux... dix-sept*, *dix-huit...*), plus ils sont grands, moins ils semblent être employés.

Parmi les adjectifs qualificatifs, certains sont beaucoup plus fréquemment employés que d'autres : *beaux* (671), *grand* (638), *dernier* (431), *folles* (321), *vieux* (282), *plein* (208), *nouvelle* (191), *ancien* (73), *moyen* (50). Nous vérifions ici ce dont nous avons parlé précédemment. Mis à part l'adjectif *dernier*, les adjectifs d'une seule syllabe sont généralement plus fréquents que ceux de deux (ou plus) syllabes.

La fréquence est aussi en rapport avec le sens des mots : plus les mots ont un sens

général, plus ils peuvent se trouver dans des contextes généraux et ainsi regrouper un plus grand nombre d'items. Par exemple, l'adjectif *beau* peut s'utiliser avec des personnes, des animaux, des végétaux, des concepts abstraits, des choses concrètes et à peu près tout ce qui nous entoure, ce qui n'est certes pas le cas d'autres mots tels que *moyen* ou *plein*.

En ce qui concerne le taux de cooccurrence, on retrouve le même phénomène que pour les enchaînements, plus les mots sont rares pris séparément, plus leur taux de cooccurrence est bas. Par contre, ce qui est un peu surprenant, c'est que parmi les liaisons dites obligatoires, la fréquence de certains groupes de mots varie énormément. Les éléments qui composent cette catégorie sont en fait trop différents pour constituer une seule et même catégorie. Cela nous pousse à en conclure qu'une telle catégorie demande à être « réévaluée », ce dont nous discuterons dans la conclusion.

Notons qu'il était difficile d'analyser la fréquence de cooccurrence de certaines séquences telles que dans l'enchaînement *grâce à*, étant donné que c'est une locution prépositive et qu'on ne peut pas vraiment considérer les mots isolément. Dans ce cas, on ne peut vraiment considérer que la cooccurrence des mots car, si les mots sont séparés, ils ont un sens différent (*la grâce*). Cela explique que pris isolément le mot soit peu fréquent, mais qu'avec la préposition il soit beaucoup plus fréquent. En analysant la fréquence de cooccurrence, on voit que plusieurs facteurs contribuent à ces résultats : fréquence, productivité et complexité des structures syntaxiques et phonétiques.

### **7.2.1. La banque de données de *Vocabprofil***

Nous avons choisi d'utiliser *Vocabprofil*, parce que c'est une banque de données différente qui nous permet de comparer les résultats obtenus avec *Lexique*. D'autre part, cette banque de données se base sur des principes différents pour établir la fréquence : la fréquence

des mots, les uns par rapport aux autres et nous permet d'avoir une vue d'ensemble sur la fréquence des mots.

Le logiciel a pourtant quelques limites. Il ne semble pas reconnaître les mots formés avec un trait d'union, ainsi que les expressions formées de plusieurs mots. Cela veut dire que *Moyen-âge* est classé au même titre que *certain âge*. Ensuite, si un mot est au 1001<sup>e</sup> rang et un autre au 1999<sup>e</sup> rang, le logiciel ne précise rien quant à leur fréquence relative puisqu'ils appartiennent tous les deux à la même catégorie K2, alors qu'un est beaucoup plus fréquent que l'autre. Cela signifie qu'on ne peut pas voir avec précision, les résultats à l'intérieur de chaque catégorie.

Toutefois, les avantages nous semblent surpasser les petits inconvénients. Les données de *Lexique 3* nous renseignent sur la fréquence de ces mots dans les films, combien de fois certains mots apparaissent par million de mots. Toutefois, cette banque de données ne nous permet pas de savoir à quelle catégorie de fréquence les mots appartiennent les uns par rapport aux autres.

Savoir qu'un mot a été utilisé plus de 1000 fois dans divers films ne nous dit pas s'il appartient aux 1000 mots les plus fréquents. C'est la raison pour laquelle nous avons choisi d'utiliser *Vocabprofil*. Afin de déterminer le rôle de la fréquence lexicale, nous avons donc attribué à chaque mot un indice correspondant aux résultats obtenus avec *Vocabprofil*, allant de 0 à 3 : K1 (1000 mots les plus fréquents) = 3. K2 (2000 mots les plus fréquents) = 2. K3 (3000 mots les plus fréquents) = 1. HL (Hors Liste) (mots n'appartenant à aucune de ces catégories) = 0.

Nous avons enlevé les mots qui n'étaient pas liants dans le texte d'origine de l'étude (236 mots). Nous avons enlevé tout ce qui n'était pas pertinent pour l'étude et n'avons gardé que les contextes de liaisons et enchaînements. Voici le texte soumis au logiciel

*Vocabprofil* pour que soient analysés tous les contextes liants du type : *Cette Anglaise\_a aux\_invités les affaires des\_étudiant.*

### **Texte de l'étude « décortiqué » en séquences liantes soumis au logiciel *Vocabprofil***

Cette Anglaise a aux invités les affaires des étudiants.  
Tes assistants, mes ouvriers et leurs enfants nos amis.  
bon ami a moyen-âge, certain âge, chantait en plein air.  
mon avis ton enfant aucun effort s'adapter à son école  
premier amour fait accompli est un léger ennui  
sont allés au dernier étage l'ancien édifice pas allés au premier  
nouveaux étudiants ont attendu second étage grand immeuble  
vieux écrits nouvelles idées belles acclamations  
Un oiseau aux yeux les oies des yogourts les yachts  
beaux allemands ont eu folles aventures grâce à faux-airs  
sous-entendu met un accent aigu d'États-Unis  
Cet homme a d'un hôpital où il y avait vieux habits et vieilles éponges  
personnes ont écrit des histoires ces hôpitaux  
Les handicaps des hollandais des hongrois des héros  
Deux amis dix ans vus à six heures trois arbres  
Il a vingt trois ans elle a vingt cinq ans neuf heures il aura cent ans  
Un héros garçon intelligent a eu cet accident affreux dans un bois immense

Nous avons analysé la fréquence des mots dans le texte de notre étude avec *Vocabprofil* (annexe 10). Après les avoir classés par ordre décroissant (si tous les chiffres étaient les mêmes, nous les avons gardés dans leur ordre d'origine) nous avons obtenu les résultats suivants :

Catégories de fréquence	nombre	%
K1 Words (1 to 1000):	133	82.10%
K2 Words (1001 to 2000):	14	8.64%
3K Words (2001 to 3000):	3	1.85%
Off-List Words:	12	7.41%

**Tableau 3.8. Fréquence des mots avec *Vocabprofil***

Nous constatons que la majorité des mots du texte sont des mots très fréquents, conformément au but fixé, afin que les apprenants L2 ne soient pas handicapés par trop de mots inconnus. *Vocabprofil* propose le « diagnostic » suivant :

K1: 0-1000 [133 ] accident affaires air airs allemands allés allés ami amis amis amour ancien anglaise ans ans ans ans arbres assistants attendu au au aucun aura aux aux avait avis beaux belles bois bon cent certain ces cet cet cette cinq dans de de dernier des des des des des des deux dix école écrit écrits effort elle en enfant enfants entendu est et et étage étage états étudiants étudiants eu eu fait faux garçon grâce grand heures heures histoires homme hôpital hôpitaux idées il il il invités le/la les les les les leurs mes met mon moyen neuf nos nouveaux nouvelles oiseau ont ont ont où pas personnes plein premier premier se/si second six son sont sous tes ton trois trois un un un un un un un vieilles vieux vieux vingt vingt vus yeux

---

K2: 1001-2000 [14] adapter âge âge aventures chantait folles héros héros immense immeuble intelligent léger ouvriers unis

---

3k [3] accent accompli édifice

---

HORS LISTE [12] acclamations affreux aigu ennui éponges habits handicaps hollandais hongrois oies yachts yogourts

---

Notons que parmi les mots n'appartenant pas à la catégorie K1, la plupart des mots ont un équivalent anglais ou au moins un mot apparenté (*cognate*) :

<b>mot français</b>	<b>mot anglais</b>
adapter	<i>adapt</i>
âge	<i>age</i>
âge	<i>age</i>
aventures	<i>adventure</i>
chantait	<i>chant</i>
folles	<i>fool</i>
héros (sing)	<i>hero</i>
héros (plur)	<i>heros</i>
immense	<i>immense</i>
intelligent	<i>intelligent</i>
unis	<i>united</i>
accent	<i>accent</i>
accompli	<i>accomplished</i>
édifice	<i>edifice</i>
acclamations	<i>acclamations</i>
éponges	<i>sponges</i>
habits	<i>habits (monk habit)</i>
handicaps	<i>handicaps</i>
hollandais	<i>Holland</i>
hongrois	<i>Hungarian</i>
yachts	<i>yachts</i>
yogourts	<i>yogurts</i>

Enfin, parmi les mots un peu plus rares, seuls 7 mots (*immeuble, léger, ouvriers, affreux, aigu, ennui, oies*) n'ont aucun équivalent en anglais. Si les connaissances des mots jouent un rôle dans la production de liaisons chez les apprenants, les séquences comprenant ces mots devraient être plus difficiles pour les apprenants et ceux-ci devraient réaliser moins de liaisons dans ce contexte. Les apprenants ne sont pas pénalisés par un texte comprenant trop de mots inconnus. Le texte a été conçu pour tester des contextes différents (h muets, h aspirés, semi-consonnes, nasales, etc.), mais de façon à ce que la plupart des mots soient connus des étudiants anglophones et bien sûr des francophones.

On voit dans le tableau en annexe 10 qu'il y a un continuum allant des séquences dans lesquelles les deux mots sont fréquents, ayant un total de 6, comme dans les séquences *nos amis, trois ans, six heures*, séquences pour lesquelles les liaisons devraient être réalisées en plus grande proportion, allant à des séquences où les deux mots sont rares, ayant un total de 1, comme dans *accent aigu*.

### **7.2.2. La fréquence des mots - discussion des données**

En ce qui concerne les enchaînements, ce sont tous des mots de même fréquence, donc la fréquence des mots ne devrait pas avoir d'incidence sur leur production, mais c'est plutôt la fréquence de la structure syntaxique qui devrait jouer un rôle majeur.

En ce qui concerne les liaisons obligatoires, la majorité des mots de cette catégorie (86/102 soit 84.31%) appartiennent à la catégorie K1. Sur 102 mots (dans les 51 séquences), 7 appartiennent à la catégorie K2 (*âge, ouvriers, Unis, immeuble, âge, folles, aventures*); 3 appartiennent à la catégorie K3 (*édifice, accompli, accent*); 6 appartiennent à la catégorie hors liste (*éponges, habits, oies, acclamations, ennui, aigu*).

Nous ne considérons pas que l'addition des facteurs soit un calcul fiable et nous n'analyserons pas les données selon ces critères ajoutés. Toutefois, nous pensons que si



plusieurs difficultés sont réunies, cela devrait limiter la production de liaisons, d'où le tableau détaillé en annexe 10, et les cas les plus « difficiles à traiter » dans le tableau 3.9.

Nous avons donné un indice (ind.) de fréquence de 0 à 3 aux mots1 et aux mots2 qui correspondent à la fréquence indiquée par Vocabprofil (K1, K2, K3, HL). Puis, nous avons ajouté l'indice des mots1 + l'indice des mots 2, donnant ainsi un indice total (allant de 1 à 6).

Les cas particuliers plus difficiles à traiter sont les suivants :

	<b>Liaisons obligatoires</b>	<b>VOC</b>	<b>VOC</b>	<b>ind.</b>	<b>ind.</b>	<b>Total</b>
LO9	moyen-âge	K1	K2	3	2	<b>5</b>
LO5	mes ouvriers	K1	K2	3	2	<b>5</b>
LO36	Etats-Unis	K1	K2	3	2	<b>5</b>
LO23	grand immeuble	K1	K2	3	2	<b>5</b>
LO10	certain âge	K1	K2	3	2	<b>5</b>
LO31	folles aventures	K2	K2	2	2	<b>4</b>
LO20	ancien édifice	K1	K3	3	1	<b>4</b>
LO17	fait accompli	K1	K3	3	1	<b>4</b>
LO39	vieilles éponges	K1	HL	3	0	<b>3</b>
LO38	vieux habits	K1	HL	3	0	<b>3</b>
LO29	les oies	K1	HL	3	0	<b>3</b>
LO26	belles acclamations	K1	HL	3	0	<b>3</b>
LO18	léger ennui	K1	HL	3	0	<b>3</b>
LO35	accent aigu	K3	HL	1	0	<b>1</b>
	<b>Liaisons facultatives</b>	<b>VOC</b>	<b>VOC</b>	<b>ind.</b>	<b>ind.</b>	<b>Total</b>
LF1	ouvriers et	K2	K1	2	3	<b>5</b>
LF2	chantait en	K2	K1	2	3	<b>5</b>
LF4	s'adapter à	K2	K1	2	3	<b>5</b>
LF15	habits et	HL	K1	0	3	<b>3</b>

**Tableau 3. 9. Séquences « difficiles à traiter » d'après la fréquence avec *Vocabprofil***

Nous pensons que les séquences inférieures à un total de 4 ont peu de chances d'être réalisées. Plus les mots sont rares, plus ils peuvent causer d'hésitations et empêcher la liaison. Mais si deux mots sont rares, il y a encore moins de chances que la liaison soit réalisée.

Les séquences potentiellement « problématiques » d'après la fréquence des mots sont : *folles\_aventures*, *léger\_ennui* et *accent\_aigu*. Les deux mots de ces séquences sont plus rares que les autres mots du texte. *Folles* et *aventures* appartiennent tous les deux à la série K2, *léger* appartient à la série K2 et *ennui* à la série hors liste; enfin *accent* appartient à la série K3 et *aigu* à la série hors liste.

Autrement dit, si l'on ne considère pas les catégories syntaxiques ou phonétiques, mais seulement la fréquence des mots, parmi les liaisons obligatoires, les séquences ayant le moins de chances d'être liées sont par ordre décroissant :

1. Accent aigu
2. Léger ennui
3. Belles acclamations
4. Vieux habits
5. Vieilles éponges
6. Folles aventures etc.

En ce qui concerne **les liaisons facultatives**, selon le même critère de fréquence des mots, les séquences pour lesquelles les liaisons ont le moins de chances d'être produites sont :

1. habits et
2. ouvriers et
3. chantait en
4. s'adapter à

Les séquences les plus susceptibles de provoquer une liaison devraient être les séquences les plus fréquentes : avec *pas* et avec les verbes ou auxiliaires *avoir* et *être*.

En ce qui concerne **les liaisons interdites**, il est intéressant de remarquer que 10 mots sur 13 sont des mots « rares », notamment les mots commençant par une semi-consonne (*yacht*) et par un h aspiré. Par contre, les mots **fréquents** commençant par une semi-consonne

(yeux) et un h muet (*hôpital, histoire*) appartenant à la catégorie K1 sont généralement classés comme « liaisons obligatoires », alors que les mots plus rares font partie des « liaisons interdites ».

Toutefois, la fréquence seule ne garantit pas forcément que la liaison sera obligatoire, ni même possible. Par exemple, la conjonction *et* est l'un des mots les plus fréquents (12 909 dans *Lexique*) de la langue française et pourtant la liaison est impossible. Il faut d'abord qu'il y ait une consonne latente pour que la liaison puisse se réaliser, ce qui n'est pas le cas ici. Et n'ayant pas de consonne latente, ce n'est pas un cas possible de liaison; le mot n'a qu'un t écrit, mais n'a pas de représentation phonologique. Pourquoi est-ce le cas ?

Tout d'abord, la structure syntaxique comprenant des conjonctions est l'une des plus rares. C'est pourquoi il n'y a aucune liaison obligatoire avec les conjonctions, sauf dans des expressions figées. D'autre part, le taux de cooccurrence est forcément bas puisque *et* n'a pas une cohésion forte avec d'autres mots, sauf dans les expressions figées. C'est justement là qu'il y a une liaison, comme dans *Pont\_ et Chaussées, Monts\_ et merveilles*, mais la liaison se fait à gauche et non à droite.

Enfin, *et* est l'homonyme de *est*. Pour éviter les « confusions » possibles, tout comme les héros/les zéros, la liaison ne se fait pas.

Bref, nous ne considérons ici pour le moment qu'un seul indice de fréquence : la fréquence des mots (mots1 et mots2). Il en existe d'autres : la fréquence de cooccurrence, la fréquence de la structure syntaxique et la fréquence des phonèmes de liaison.

### **7.2.3. La fréquence des mots1 et des mots2 de notre texte et le codage**

Nous avons analysé la fréquence des mots1 et des mots2 en utilisant les banques de données de *Lexique* et de *Vocabprofil* et avons ainsi obtenu les résultats dans le tableau en annexe 11. Nous avons codé les occurrences et toutes les informations pertinentes avec des

lettres, des nombres et autres signes diacritiques, pour que ces renseignements puissent être analysés par le logiciel *GoldVarb* (de Rand et Sankoff, 1990).

Pour que le logiciel puisse analyser les données, il fallait donner aux renseignements que nous voulions analyser des codages différents. C'est pourquoi on ne pouvait pas utiliser les mêmes lettres ou signes diacritiques. Par exemple, pour analyser la fréquence des mots avec *Lexique* nous avons utilisé des lettres minuscules, pour la fréquence de cooccurrence, des lettres majuscules, pour la fréquence avec *Vocabprofil*, des chiffres, etc.

Nous avons calculé les liaisons réalisées et le type de réalisation (enchaînée, non-enchaînée, avec reprise, etc.). Les occurrences de liaisons et enchaînement étaient la variable stable. Les facteurs linguistiques tels que le type de structure syntaxique, le phonème de liaison, la fréquence des mots et de leur cooccurrence et les facteurs sociaux propres aux locuteurs (origine géographique, contact avec le français, etc.) sont tous présents dans l'analyse des données.

Nous avons d'abord codé les liaisons « correctement » réalisées dès la première tentative (a) et les avons considérées comme telles, seulement si elles ont été enchaînées, si la consonne de liaison a été bien prononcée, si le phonème du Mot2 suivant la consonne de liaison a été « correctement prononcé » et si la séquence a été réalisée selon les normes orthoépiques du français standard.

Nous avons attribué lors du codage une lettre qui correspond aux situations suivantes :

- a = bien réalisée, bien enchaînée (dès la première tentative)
- b = non réalisée
- c = consonne liaison produite, mais non-enchaînée
- d = consonne liaison mal prononcée
- e = reprise avec bonne liaison
- f = reprise sans liaison
- g = reprise avec mauvaise liaison
- h = fausse liaison (pataquès): consonne rajoutée pour faire une liaison
- i = fausse liaison (pataquès) + reprise + correction (et bonne liaison)
- j = phonème ou syllabe du mot d'avant mal prononcée (par exemple, *vieille*)
- k = phonème ou syllabe du mot d'après mal prononcée
- l = fausse liaison (pataquès) + reprise PAS DE liaison

Par exemple, si un sujet a prononcé /d/ dans la séquence *grand immeuble* au lieu de /t/, nous l'avons codé « d » (consonne liaison mal prononcée). De même, si un sujet a prononcé le h de *héros*, nous l'avons codé « k » (phonème ou syllabe du mot d'après mal prononcée) puisque, d'après les normes du français standard, on ne prononce jamais le h. Les chiffres correspondant aux liaisons réalisées que nous verrons dans le chapitre suivant sont ceux codés « a » (bien réalisée, bien enchaînée).

## 8. La fréquence et la productivité des structures syntaxiques

Les études sur les liaisons, des plus anciennes aux plus récentes, prennent en compte les structures syntaxiques. C'est le cas de l'Académie Française, Delattre (1951), Fouché (1959), Grevisse (1980) ou Boula de Mareüil, Adda-Decker et Gendner (2003). En effet, on ne peut pas considérer que les liaisons sont un phénomène uniquement phonétique. Si tel était le cas, toutes les liaisons seraient traitées de façon identique dans un même contexte phonétique, et nous savons bien que ce n'est pas le cas.

Les structures syntaxiques jouent aussi un rôle majeur dans la production de liaisons. Selon Ellis (2002 : 148), les catégorisations humaines seraient plus sensibles à la fréquence des structures qu'à la fréquence des mots. De même, selon Bybee (1995) et Bybee et Thompson (2000), la productivité des modèles phonologiques, syntaxiques et morphologiques dépendrait plus de la fréquence des constructions que de la fréquence des mots.

La fréquence des constructions en détermine leur productivité parce que

- (a) plus des items lexicaux ont été entendus dans une certaine position pour une construction, moins il y a de chances que cette construction soit associée à un item lexical en particulier et plus il y a de chances qu'une catégorie générale soit formée à partir des items qui apparaissent dans cette position ;

- (b) plus une catégorie contient d'items, plus ses traits sont généraux et plus il est possible que de nouveaux items puissent s'inscrire dans cette classe ;
- (c) une haute fréquence des structures assure qu'une construction soit utilisée fréquemment renforçant ainsi le schéma représentationnel, rendant plus accessibles les nouveaux items (Bybee et Thompson, 2000).

Nous souhaitons analyser le rôle de la fréquence des structures syntaxiques pour déterminer si le degré d'étroitesse dont parlait Delattre (1955 : 45-46) est un facteur significatif dans la production de liaisons et dans quelle mesure. Celui-ci a établi une hiérarchie du degré d'étroitesse entre les mots, qui correspond à la fréquence de cooccurrence de ces classes de mots allant de 1 à 10 (10 étant les liaisons les plus fréquemment produites et 1 les moins fréquemment produites), dont nous avons parlé au chapitre 1. Pour Delattre la fréquence d'occurrence de ces liaisons correspond à leur « degré d'union ».

Nous apporterons quelques précisions à cette classification et y ajouterons notre indice de fréquence allant de  $a$  à  $m$ ,  $a$  étant la structure la moins fréquente et  $m$  la structure la plus fréquente.

Delattre (1956) propose une liste des différents niveaux de fréquence : liaisons très fréquentes, assez fréquentes, mi-fréquentes, peu fréquentes, rares, très rares, et il propose ici un continuum de liaisons, modèle légèrement différent de la traditionnelle tripartition : obligatoires, facultatives, interdites. Il précise ainsi la typologie abordée en 1955.

Nous suivrons les catégorisations traditionnelles pour établir une hiérarchie des séquences de notre texte. Nous analyserons les résultats obtenus avec *GoldVarb* pour déterminer le rôle des structures syntaxiques. Notre hypothèse est que les structures syntaxiques les plus fréquentes telles que [dét. + nom], auront un taux de liaison plus élevé que les liaisons les moins fréquentes telles que les expressions figées.

La fréquence des structures syntaxiques devrait être au moins aussi importante que la fréquence des mots pour les raisons suivantes. Plus les structures sont fréquentes, plus elles sont entendues et utilisées et plus elles ont de chances d'être stockées dans le lexique et d'entraîner une liaison. Si le contexte [adj. poss. + nom] est plus fréquent que le contexte [adj. indéf. + nom], un pourcentage plus élevé de liaisons devrait être effectué dans ce contexte.

C'est d'ailleurs ce qui pourrait expliquer que dans les textes à lire de Mastromonaco, même si les mots<sup>2</sup> sont de différente fréquence, les différences de production entre la lecture des deux textes (des deux groupes) ne sont pas significatives. Si les structures syntaxiques sont de même fréquence, alors, la fréquence lexicale ne jouerait pas un rôle significatif.

Par contre, des différences de fréquence dans les structures syntaxiques devraient avoir des effets sur la production de liaisons, notamment avec les liaisons facultatives. Le contexte [aux. + part. passé] paraît plus fréquent que le contexte [verbe + prép.]. Ainsi, le contexte [aux. + part. passé] devrait engendrer plus de liaisons que le contexte [verbe + prép.].

La structure la plus fréquente est la structure [dét. + nom]. Or, une grande partie des liaisons interdites se trouvent dans ce contexte [dét. + nom] avec des h aspirés, des semi-consonnes et d'autres cas particuliers. En revanche dans la grande majorité de ces contextes syntaxiques, la liaison est produite de façon systématique (*les enfants, un arbre, etc.*).

Les cas de liaisons interdites sont donc des exceptions à l'intérieur de cette structure. Pour cette raison, nous pensons que, si des erreurs apparaissent, elles devraient consister à généraliser la règle, à faire des liaisons en situation de liaison interdite dans le contexte [dét. + nom]. Autrement dit, si une partie des liaisons interdites est produite, cela devrait arriver dans des structures les plus fréquentes de ce type.

Ainsi, d'après la loi de la hiérarchie et du nivellement, les liaisons interdites avec les h aspirés pourraient se régulariser et les liaisons se réaliser. C'est en effet le cas avec par exemple le mot *haricot*, cas pour lequel il y a eu un vif débat dans lequel il était question

d'accepter la liaison, tant elle était produite par un grand nombre de personnes. La réponse finale a été de conserver les règles du h aspiré (pas de liaison). Il n'en demeure pas moins que de nombreux francophones réalisent la liaison, que cela soit conforme à la norme prescrite ou non.

D'un autre côté, la fréquence de certains mots peut avoir l'effet contraire : comme ces mots sont souvent entendus, la fréquence renforce la règle qui consiste à ne pas produire de liaison. Les locuteurs ont acquis dans leur lexique les séquences les plus fréquentes qu'ils ont entendues dans leur entourage. De plus, cet effet peut être renforcé si la production de liaisons engendre une ambiguïté.

Par exemple, si la séquence *les héros* contient une liaison, elle pourrait être « confondue » avec *les zéros* (les deux séquences étant alors des homonymes), de même que la séquence *les hauteurs* pourrait être « confondue » avec *les auteurs*. Quand nous disons « confondue », nous voulons dire que dans l'esprit des locuteurs francophones, la prononciation de liaison avec les h aspirés pourrait évoquer les images des homonymes correspondants. Ainsi, pour ne pas prononcer *les héros* comme *les zéros*, les locuteurs marqueraient intuitivement la différence des séquences potentiellement homonymes, en ne faisant pas la liaison.

Ce qui fait alors la différence entre le renforcement de la règle pour les mots les plus fréquents (donc le maintien de la non-liaison) et au contraire le nivellement, c'est-à-dire la production de la liaison, ce serait plutôt combien de fois les locuteurs ont entendu ces mots, et combien de fois ils les ont entendus avec ou sans liaisons, ainsi que le nombre d'items appartenant à une structure. Par exemple, sur *Lexique*, on voit que *héros* (75) et *haricot* (13) sont plus fréquents que *hâche* (10), *handicap* (3) et *hampe* (0). Mais comme on ne peut pas faire la liaison avec *les héros*, (qu'on prononcerait alors /zeRo/) cela l'exclut de la liste des séquences liantes possibles. Donc, l'un des mots les plus fréquents parmi les h aspirés est le



mot *haricot*. Ainsi, si un locuteur a entendu dans son entourage cette séquence avec une liaison, il l'aura sans doute entendu suffisamment de fois pour que celle-ci soit acquise dans son lexique comme appartenant à la catégorie des séquences avec h muets (avec liaison).

La plupart des mots commençant par un h (que ce soit un h muet ou un h aspiré) appartiennent à la catégorie hors liste (avec h muets : *hôpital, huitre, huissier, hippodrome* et avec h aspirés : *hache, héron, hamac, hangar*), leur taux de fréquence est bas.

Les mots commençant par un h aspiré constituent « une classe à part », car ils sont en majorité (et encore plus que les mots commençant par un h muet) des mots « rares » (appartenant à la catégorie K3 ou Hors Liste). Pour confirmer notre hypothèse sur le rôle de la fréquence des mots dans la prononciation des liaisons, nous avons testé la fréquence de 54 mots que nous estimons être les plus fréquents commençant par un h aspiré, dont des listes plus exhaustives (mots plus rares tels que *hâbleur, hachure, haddock, halage...*) se trouvent dans les manuels de français à l'usage des francophones, notamment dans Grevisse (1980).

Ainsi, selon l'analyse de *Vocabprofil*, sur une liste de 54 mots les plus « fréquents » commençant par un h aspiré) nous obtenons les catégorisations suivantes :

- 1 mot appartient à la série K1 :

*haut*

- 5 mots appartiennent à la catégorie K2 :

*hasard, hausse, hauteur, héroïne, héros*

- 3 mots appartiennent à la catégorie K3 :

*hiérarchie, honte, hurler*

- 45 mots appartiennent à la catégorie hors liste :

*hache, haie, haillons, haine haïr hale, hall, halo, hamac, hameau, hanche, handicap, hangar, hanter, happer, harasser, harceler, hardi, harem, hareng, haricot, harpe, harpon, haschisch, hâte, hérissier, hernie, héron, hêtre, heurter, hibou, hideux, hocher, hockey, hollandais, homard, hongrois, honteux, hoquet, hotte, housse, hublot, huer, huitième, hurlement.*

Ainsi, si on ne fait pas de liaison avec les mots commençant par un h aspiré, ce n'est pas parce qu'ils sont d'origine germanique (ou autre que grecque et latine), mais c'est parce que les mots de ces origines sont moins fréquents que les mots d'origine grecque et latine, qu'ils ont évolué en une classe à part avec laquelle on ne fait pas la liaison. En d'autres termes, ce n'est pas parce que ces mots ont un h muet qu'on ne fait pas la liaison, mais c'est

parce que ces mots sont moins fréquents qu'ils ont une prononciation différente des autres (donc un h aspiré).

Nous voyons ici l'effet de la fréquence sur l'évolution de la langue et la prononciation des mots. Étant donné que la consonne finale continue à être prononcée uniquement pour les mots les plus fréquents, on ne fait pas de liaison avec les mots les plus rares (tels que ceux commençant par un h aspiré ou une semi-consonne). Nous voyons aussi que l'effet de fréquence de mots isolés, contribue à créer une nouvelle catégorie de mots, si bien que fréquence de mot (*token*) et fréquence de classe (*type*) tendent à se rejoindre.

Non seulement ces mots sont rares, mais ils sont aussi, avec les mots commençant par une semi-consonne, des exceptions, puisque ce n'est qu'avec ces mots qu'on ne fait pas de liaison dans un contexte pourtant obligatoire [dét. + nom]. Rien ne permet aux francophones de déterminer s'il faut faire la liaison ou non, si ce n'est qu'ils ont entendu les séquences et ont peut-être été corrigés à la maison ou ailleurs. Après avoir entendu les séquences un certain nombre de fois étant enfants, leur intuition leur dicte s'il faut faire la liaison ou non. L'intuition consiste à récupérer les formes stockées dans le cerveau et à puiser dans des catégories. Toutefois, si ces catégories ne sont pas clairement établies et délimitées, les erreurs semblent inévitables.

Il est possible que les francophones (et en particulier les francophones minoritaires) et les anglophones traitent les h aspirés ou au moins certains mots ayant un h aspiré comme les autres mots du même contexte [dét. + nom], s'ils les ont entendu un certain nombre de fois pour qu'ils soient assez familiers, mais pas assez souvent dans certains contextes (par exemple au pluriel), pour avoir enregistré les séquences comme étant sans liaison. C'est-à-dire qu'ils ne les auraient pas enregistrées dans leur lexique comme appartenant à une catégorie à part. Il est alors possible que les liaisons soient réalisées avec les mots les plus fréquents

contenant un h aspiré, par un processus de nivellement, de régularisation et de surgénéralisation.

Comme les mots *haut* (+ mots de la même famille) et *héros* (au singulier comme au pluriel) sont les mots les plus fréquents parmi les mots commençant par un h aspiré, nous formulons l'hypothèse que, si des participants font une liaison avec les h aspirés, ce sera dans ces contextes où ils sont les plus fréquents de leur catégorie. D'un autre côté, si les locuteurs ont entendu suffisamment de fois certaines séquences contenant des h aspirés sans liaison ni élision, il est également possible qu'ils aient enregistré ces séquences dans leur lexique et qu'ils ne fassent donc pas la liaison.

La liaison avec *les héros* étant ambiguë, les locuteurs pourraient faire la liaison au singulier (et aussi par analogie avec le féminin, puisqu'on fait un enchaînement avec *une héroïne*, et une liaison avec *des\_héroïnes*), sans faire la liaison au pluriel puisqu'ils prononceraient ainsi *les héros* comme *les zéros*. Les locuteurs pourraient donc faire des liaisons « fautives » avec les mots les plus fréquents et ne pas en réaliser avec les mots plus rares. Ils pourraient avoir créé une catégorie [h aspiré + 0 liaison], avec les mots plus rares (*hâche, handicap...*), mais les mots les plus fréquents pourraient avoir été placés (à tort) dans la catégorie h muet, qui contient généralement des mots plus fréquents (*homme, hôpital, habit, etc.*).

Autrement dit, pour analyser les structures syntaxiques, il faut considérer à la fois leur fréquence, mais aussi leurs règles, leur complexité et leurs exceptions. Nous supposons que les liaisons seront produites en plus grand nombre lorsqu'elles sont très fréquentes, mais aussi lorsqu'elles regroupent un grand nombre d'items. Par exemple la catégorie [art. + nom] est non seulement plus fréquente que la catégorie [adj. + nom], mais elle regroupe aussi un bien plus grand nombre d'items.

Le nombre de noms possibles après les articles et le nombre de séquences possibles avec les articles est presque illimité, tandis que le nombre d'adjectifs placés avant le nom et le nombre de séquences liantes possibles dans ce contexte est très limité en dehors des cas d'adjectifs pluriels. D'après le MBU, les liaisons devraient être produites en plus grandes proportions dans des cas où le cerveau a stocké des catégories contenant un grand nombre de séquences (structures productives) et devraient être réalisées en plus faibles proportions dans des catégories plus restreintes. La catégorie [art. + nom] devrait donc être la catégorie dans laquelle les liaisons sont réalisées en plus grand nombre.

## 9. La fréquence des phonèmes de liaison

Après avoir analysé la fréquence d'un point de vue lexical et syntaxique, nous allons analyser la fréquence d'un point de vue phonétique selon les consonnes de liaisons.

### 9.1. Le nombre de structures de chacune des consonnes de liaison

Pour le nombre de structures qu'englobent les consonnes de liaisons nous nous basons d'abord sur une analyse syntaxique.

Voyons d'abord les catégories que chaque consonne de liaison regroupe :

- /z/ :
- mots se terminant par s, x ou z
  - tous les déterminants au pluriel : *les, des, ces, mes, tes, ses, aux...*
  - adjectifs numéraux et cardinaux : *deux, trois, six, dix.*
  - adjectifs qualificatifs au masculin singulier, masculin pluriel et au féminin pluriel + noms : *gros appétit, bons amis, folles aventures...*
  - autres adjectifs + noms : *toutes, certains, quelques, autres, plusieurs*
  - noms au pluriel + adjectifs : *enfants intelligents...*
  - pronoms personnels sujets, COD, COI et autres : *ils, , les, nous, uns, unes, ceux...*
  - formes verbales des 1<sup>e</sup> et 2<sup>e</sup> personnes du singulier et du pluriel + ... (préposition, adverbe, etc.) : *suis en, es au, sommes encore, allez au ;*
  - auxiliaires + participes passés : *je suis arrivé, es, sommes, êtes, as, avons, avez*
  - semi-auxiliaires + formes verbales : *je vais aller, nous pourrons aller...*
  - expressions figées : *monts et merveilles, vis-à-vis ...*
  - prépositions : *dans, sans, sous, chez...*
  - adverbes monosyllabiques : *plus, très, moins, pas...*
  - s d'épenthèse : *vas-y...*

- /t/ :
- mots se terminant par t ou d
  - formes verbales se terminant par t ou d : 3<sup>es</sup> personnes du singulier et pluriel, de presque tous les modes et temps : indicatif, gérondif, subjonctif, conditionnel.
  - *c'est* + ... (n'importe quoi)
  - adjectifs fréquents se terminant par t ou d + nom : *second étage, grand enfant...*
  - verbes avec inversion se terminant par t ou d + pronom : *vend-il, sait-il...*
  - auxiliaires + participes passés : *est arrivé, sont, ont, peut, sait, peuvent...*
  - semi-auxiliaires + formes verbales : *il veut aller, ils vont aller*
  - pronoms : *tout, dont*
  - noms + adj dans expressions figées : *fait accompli, accent aigu...*
  - adverbes + verbes ou adjectifs : *vraiment utile, vraiment intelligent...*
  - expressions figées : *petit à petit, comment allez-vous, pied-à-terre...*
  - t d'épenthèse : *va-t-il, montre-t-il*
- /n/ :
- articles singuliers indéfinis : *un*
  - prépositions : *en*
  - adverbes monosyllabiques : *bien*
  - adjectifs possessifs masculins singuliers : *mon, ton, son*
  - adjectifs qualificatifs et indéfinis courts se plaçant avant le nom : *bon, ancien, aucun*
  - pronoms indéfinis : *rien, aucun, on*
- /R/
- adjectifs qualificatifs courts antéposé se terminant par un R : *premier, léger...*
  - verbes à l'infinitif : *s'adapter à, aller à ...*
- /p/
- adverbes : *trop, beaucoup*

Bien que ces catégories ne soient pas forcément complètes, car il existe de nombreux autres contextes syntaxiques, nous les avons mentionnées surtout pour expliquer pourquoi certaines liaisons pourraient être produites plus que d'autres.

Mallet (2008) a calculé les contextes de liaisons et elle est arrivée aux résultats que l'on voit dans le tableau 3.10, résultats qui correspondent à notre liste de contextes possibles d'après le contexte phonético-syntaxique. Ces résultats sont classés par ordre décroissant de contextes de liaisons. On remarque que les contextes les plus fréquents sont d'abord /z/, puis /t/ et /n/, mais le nombre de liaisons réalisées est par ordre décroissant /z/, /n/ et /t/.

Consonne de liaison	nombre de contextes de liaison	nombre de liaisons réalisées
/z/	<b>13852</b>	5960
/t/	<b>8113</b>	1888
/n/	<b>5512</b>	4970

Tableau 3.10. Contextes liants et liaisons réalisées selon les phonèmes. Mallet (2008 : 218)

Adda-Decker, Boula de Mareüil et Lamel (1999 : 2240) : arrivent à des résultats similaires avec la consonne /z/ : « 25% of the words of the lexicon used in this study have a /z/ liaison. » D'après ce que regroupent ces catégories et les sites de liaisons potentiels, les consonnes de liaisons les plus fréquentes sont les suivantes par ordre décroissant : /z/, /t/, /n/, /R/, /p/, /k/. Il en est de même pour Boula de Mareüil et Adda-Decker (2003 : 1329) qui, à partir d'un corpus de 100 heures de texte lu, ont conclu que les liaisons les plus réalisées sont par ordre décroissant de fréquence: /z/, /t/, /n/, /R/, /p/, en accord avec Léon (1992 : 152).

Nous souhaitons tester l'hypothèse selon laquelle la consonne de liaison jouerait un rôle dans la production de liaisons.

## 9.2. Consonne de liaison et structure morphosyntaxique et sémantique

Les consonnes de liaisons dépendent en partie des structures morphosyntaxiques. Parmi les adjectifs, seules certaines structures permettent l'utilisation de la consonne de liaison /z/. C'est majoritairement le pluriel de ces adjectifs, car les seuls adjectifs au singulier régulièrement antéposés au nom qui ont cette consonne sont *gros* et *vieux*.

Si l'on considère l'ensemble des séquences contenant un adjectif, on voit que les cas de liaisons sont rares. Certains adjectifs qualificatifs courts et fréquents (*beau, bon, grand, gros, petit, vieux*) sont antéposés au nom. Toutefois, la plupart des adjectifs se placent après le nom. Même les adjectifs qui peuvent être placés avant le nom sont en fait souvent placés après le nom. Étant donné qu'on ne fait pas de liaison avec certaines consonnes, cela exclut par exemple *long*. Enfin, parmi les quelques adjectifs antéposés, les contextes de liaisons sont en réalité limités, car de nombreux noms suivant ces adjectifs commencent par une consonne.

D'autre part, parmi ces adjectifs antéposés, se trouvent les adjectifs qui peuvent être placés avant ou après le nom, mais dont le sens change selon leur place : (*grand, ancien*, etc.). Cela a pour résultat que les locuteurs les ont entendus dans les deux positions et avec des sens

différents, ce qui en complique et rallonge le traitement.

Enfin, apparaît la dernière catégorie : les adjectifs qui sont le plus souvent après le nom, ou en position d'attribut, mais qui peuvent quand même être placés avant le nom. Il est possible de placer n'importe quel adjectif avant le nom, surtout pour mettre l'emphase sur l'adjectif plutôt que le nom. Par exemple, on dirait normalement : *des histoires formidables*.

Or, il est possible de dire : *de formidables histoires*, bien que ce soit plus rare. Dans ce cas, la liaison aurait moins de chances d'être produite parce que ce n'est pas la place habituelle de cet adjectif et parce que cette séquence est rare. Ceci étant, il y a quand même des cas où l'adjectif se place plutôt avant le nom, bien qu'il soit possible de le placer après, comme dans *horribles histoires* (plutôt que *des histoires horribles*).

Ainsi, lorsqu'un locuteur se trouve face à un adjectif placé avant le nom (lors d'un exercice de lecture par exemple), il est confronté à un triple dilemme. D'une part, il y a la contrainte anti-hiatus qui l'incite à lier les mots. D'autre part, la plupart des adjectifs (courts et fréquents antéposés) dans cette position prennent normalement une liaison. Mais il y a l'effet inverse : le taux de cooccurrence de ces deux mots est faible, ces séquences ne sont pas stockées dans le lexique du locuteur ce qui a pour effet de traiter les mots isolément et donc de ne pas réaliser la liaison. Cela pourrait créer des hésitations et des reprises qui se solderaient pas une liaison non réalisée.

Prenons comme exemple les séquences de notre texte *léger ennui* et *second étage*. Il suffit de consulter le dictionnaire pour constater que la plupart du temps, l'adjectif *léger* se place après le nom. On dirait plutôt *une tenue légère, un sommeil léger, des mœurs légères*, etc., mais dans d'autres expressions, il se place plutôt avant (*léger accident, léger ennui*). De plus, les mots *léger* et *ennui* sont tous les deux des mots peu fréquents. D'après *Vocabprofil*, *léger* appartient à la catégorie K2 et *ennui* à la catégorie Hors-Liste. Enfin, la consonne de liaison /R/ est une consonne de liaison plus rarement produite. Cette séquence aurait donc peu

de chances d'être produite, même dans les expressions où l'adjectif se place avant, car le taux de cooccurrence est faible, les mots<sup>1</sup> et les mots<sup>2</sup> sont peu fréquents, la structure syntaxique est peu fréquente et peu productive, et la consonne de liaison est rare (R).

En ce qui concerne le mot *second*, le *Petit Robert* (2002, p. 2389) propose la définition suivante: « on emploie plutôt *second* quand il n'y a que deux choses » sinon, s'il y a plus que deux choses, on emploie plutôt le mot *deuxième*. Donc, la séquence *second étage (du grand immeuble)* non seulement n'a pas un taux de cooccurrence élevé puisqu'il est plutôt rare de trouver des maisons ou des immeubles de deux étages, mais le contexte de la phrase dit précisément le contraire. En général, un immeuble a plusieurs étages.

Le locuteur se trouve encore face à un dilemme. Par analogie à d'autres adjectifs (Notamment les autres adjectifs ordinaux : *premier*, *deuxième*, *troisième*, etc.) et par rapport aux adjectifs tels que *grand* qui, lui, est souvent placé avant le nom, la tendance serait de faire une liaison et de prononcer d /t/. Toutefois, comme la position antéposée au nom de *second* n'est pas très fréquente, et elle est encore moins fréquente dans la séquence *second étage*, cette séquence n'a probablement pas été stockée dans le lexique des locuteurs. Cela devrait avoir comme conséquence possible, soit que la liaison ne soit pas effectuée, soit que le d soit prononcé /d/ au lieu de /t/ ou que la liaison soit d'abord non-enchaînée, qu'il y ait reprise suivie d'une liaison non réalisée. Une telle séquence devrait aboutir à un faible taux de réalisation et à des erreurs.

Toutefois, étant donné que les deux mots de cette séquence (*second* et *étage*) sont tous les deux des mots très fréquents (K1), et que la consonne de liaison est plus fréquemment réalisée en /t/ qu'en /R/, nous pensons que le taux de réalisation de la liaison de cette séquence devrait être plus élevé que la séquence *léger ennui*, mais devrait tout de même être plus bas que d'autres séquences de la même catégorie syntaxique de type *vieux écrits*.

Cette problématique des adjectifs antéposés explique d'ailleurs pourquoi certains



adjectifs se terminant par une nasale sont dénasalisés (*Moyen-Âge*) alors que d'autres ne le sont pas, et pourquoi dans une même position, la liaison n'est pas faite du tout (voir Sampson, 2001). La production de liaisons dépend de la fréquence de ces adjectifs en général, de la fréquence de cooccurrence entre le mot1 et le mot2 et surtout de la fréquence de l'adjectif dans cette position. La consonne de liaison joue aussi un rôle.

Dans l'ensemble, nous pensons que mis à part les cas particuliers des adjectifs au singulier, les phonèmes /z/ et /t/ regroupent le plus grand nombre de structures et sont donc plus productives. C'est la raison pour laquelle les erreurs de liaisons consistent le plus souvent à insérer un /z/ ou un /t/. Si la consonne de liaison joue un rôle significatif, c'est avec ces deux phonèmes qu'on devrait trouver des différences par rapport aux autres consonnes de liaison.

### 9.3. La fréquence des consonnes de liaisons

Pour déterminer la fréquence des consonnes de liaisons, il faut prendre en considération non seulement le nombre de sites potentiels, mais aussi le nombre de liaisons réalisées d'après, entre autres, la fréquence des items de ces structures. Par exemple, si 90.2 % des sites potentiels de liaisons en /n/ (5512) sont réalisés alors que seulement 23.3 % des sites potentiels de liaisons en /t/ (8113) sont réalisés, au final, une plus grande proportion de liaisons en /n/ sera réalisée, bien qu'au départ les sites potentiels étaient plus réduits avec cette consonne de liaison.

À titre informatif, nous avons trouvé la fréquence des articles sur *Lexique 3*:

	Fréquence films	Fréquence livres
<b>singulier</b>		
la	14946	23633
le	13652	18310
<b>un</b>	<b>12087</b>	<b>13550</b>
une	7907	9587
<b>pluriel</b>		
les	8720	14662
des	6055	10624
aux	835	1907

**Tableau 3.11. Fréquence des articles dans les films et livres avec *Lexique***

Nous remarquons que les formes au singulier sont plus fréquentes que les formes au pluriel. Par contre, les formes liantes sont aussi plus restreintes au singulier. *Le, la* et *une* ne permettent pas de liaison (*une* permet un enchaînement). Parmi les articles au singulier, seul *un* permet une liaison. Celui-ci est, d'après la fréquence dans les films, le plus fréquent de tous les articles, il est deux fois plus fréquent à l'oral que *des*.

La liaison en /n/ en général ne comprend qu'une classe très restreinte d'items :

- Déterminants : un, mon, ton, son...
- Quelques rares adjectifs : ancien, bon...
- L'adverbe : bien

Mais ces items sont très fréquents et même les plus fréquents de leur catégorie syntaxique. C'est la raison pour laquelle selon Mallet (2008 : 256), la hiérarchie des pourcentages de réalisation des consonnes de liaison est la suivante :

/n/	/z/	/t/
96,6 % LR	59,1 % LR	32,4 % LR

**Tableau 3.12. Distribution des consonnes de liaison réalisées (LR) /n/, /z/ et /t/. Mallet (2008 : 256)**

Et le total de liaisons réalisées est le suivant :

Consonnes de liaison	Contextes de liaisons possibles	Liaisons réalisées		Liaisons non réalisées	
		Nombre total de liaisons réalisé	% de liaisons réalisées	Nombre total de liaisons non réalisées	% de liaisons non réalisées
/z/	13 852	5 960	43 %	7 892	57 %
/n/	5 512	4 970	90,2 %	542	9,8 %
/t/	8 113	1 888	23,3 %	6 225	76,7 %
/R/	942	12	1,3 %	930	98,7 %
/p/	98	11	11,2 %	87	88,8 %
/k/	2	0	0 %	2	100 %
<b>Total</b>	<b>28 519</b>	<b>12 841</b>	<b>45 %</b>	<b>15 678</b>	<b>54 %</b>

**Tableau 3.13. Fréquence des liaisons réalisées et non réalisées. Mallet (2008 : 213)**

En nous inspirant de la typologie de Delattre (1955 : 45-46), nous pensons que les structures syntaxiques les plus fréquentes (par rapport aux autres structures syntaxiques) sont les suivantes (tableau 3.14.), classées par ordre décroissant de fréquence :

Structure	Phon.	exemple
1. Articles + noms	/z/ /n/	Les amis, des amis Un ami
2. Adjectifs poss, dém, inter, num, card... (determinants autres qu'articles + noms)	/z/ /n/	Mes, tes, ses, nos, vos, premiers... Mon, ton, son
3. Pronom : pronom clitique + verbe Verbe + pronom	/z/ /t/	Ils ont, il nous aime, je vous aime Va-t-il, vient-il
4. Préposition monosyllabique	/z/ /n/	Dans une heure, chez elle En,
5. Adverbes monosyllabiques	/z/	Très, pas
6. Conjonctions monosyllabiques	/t/	Quand il a eu
7. C'est + ...	/t/	C'est utile
8. Auxiliaires + participe passés	/t/	Est arrivé, sont arrivés
9. Adjectifs qualificatifs + noms	/z/ /n/	Beaux enfants Ancien immeuble
10. Certaines expressions figées	/z/ /t/ /n/	États-Unis Fait accompli Moyen-Âge
11. Formes verbales	/t/	Veut arriver, faut aller, arrivent à
12. Adjectifs antéposés au nom Verbes à l'infinitif	/R/ /R/	Premier, dernier, léger Aller en ville

**Tableau 3.14. Fréquence des phonèmes de liaisons d'après les catégories syntaxiques**

Il existe certainement d'autres catégories qui ne sont pas dans ce tableau (3.14.).

Celui-ci ne se veut pas exhaustif, il existe bien sûr d'autres catégories (prépositions et adverbes polysyllabiques, etc.). Il n'est présent qu'à titre indicatif, afin d'expliquer notre raisonnement. Certaines structures syntaxiques sont plus fréquentes et plus productives que d'autres. Selon cette fréquence syntaxique et le nombre d'items qu'une structure englobe, cela a un impact sur les consonnes de liaisons. Par exemple, si une consonne de liaison ne se trouve que dans des structures plus rares et moins productives, comme le /R/ qui ne se trouve qu'avec quelques adjectifs antéposés et les verbes à l'infinitif + préposition ou adverbe, le nombre de contextes possibles est très restreint, donc la consonne de liaison /R/ devient une consonne peu productive. Ainsi, la liaison avec /R/ se fait le plus souvent dans un contexte très formel et elle est rarement réalisée.

On fait le plus souvent une liaison avec les phonèmes /z/, /n/ et /t/, parce que ce sont les consonnes les plus fréquentes dans les mots les plus fréquents. La liaison avec /p/ se limite

à deux items : *trop* et *beaucoup*. De plus, les liaisons avec ces deux items étaient autrefois facultatives fréquentes, mais elles sont maintenant plutôt rares et ne s'emploient que dans des situations formelles. Ceci est d'ailleurs dû au manque de productivité de ces consonnes.

On fait le plus souvent la liaison entre un mot plus court et un mot plus long (d'où la liste d'adjectifs antéposés courts et fréquents dans les manuels de grammaire du français FLE/FLS). Or, le mot *beaucoup* ayant deux syllabes, cela réduit encore les occurrences de cette liaison. Si celle-ci diminue, cela diminue en même temps l'effet d'analogie pour *trop*, ce qui, à long terme prédit la disparition des deux.

De plus, Thomas (1998 : 544) explique avec humour que la liaison avec /p/, peut être pour le moins ambiguë : « Si vous avez été trop souvent à Montréal, évitez de dire que vous avez trop\_été au Québec. » Dans le même ordre d'esprit, il vaudrait mieux éviter aussi de dire « qu'il est trop\_aidé ». Ces cas d'homophonie rajoutent peut-être aux hésitations possibles des locuteurs quant à faire une liaison avec ces mots.

Parce que la structure est également peu productive avec le g écrit, puisqu'il y a très peu d'items dans ces catégories phonétiques, la consonne de liaison avec /k/ a disparu alors qu'elle était encore utilisée il y a quelques décennies avec *sang impur*. Le locuteur n'ayant pas assez d'items dans la catégorie [adj.-g + nom], ne fera pas la liaison avec le mot *long*, comme dans *long hiver*.

Il ressort de tout ceci, que le phonème /z/ regroupe le plus de structures fréquentes suivi du /n/ suivi du /t/. Même si /t/ est légèrement plus fréquent que /n/, /t/ comprend un nombre d'items plus restreint (*quand, est*) et parmi certaines de ces structures, la liaison n'est pas obligatoire (avec formes verbales, avec *quand, c'est* et *est*). Le phonème /n/ contient un peu moins de structures fréquentes (déterminant, préposition monosyllabique, expressions figées), mais la structure [dét. + nom] est la plus fréquente et dans la plupart des cas la liaison est obligatoire (sauf entre un adjectif et un nom).

Durand et Lyche (2008 : 58) arrivent à des résultats similaires en précisant que le nombre global de liaisons sur 2261 sites potentiels pour chaque consonne est le suivant :

[z]: 4544 > [n]: 3689 > [t]: 1665 > [r]: 13 > [p]: 9. Nous pensons retrouver des similarités dans le taux de réalisation des liaisons de notre étude.

#### **9.4. Les contextes phonétiques de la consonne de liaison**

Nous allons étudier le traitement des différentes consonnes de liaison. Nous reconnaissons aussi que d'autres facteurs phonétiques semblent jouer un rôle sur la production de liaisons, notamment : les phonèmes se trouvant avant la consonne de liaison et le nombre de liaisons et enchaînements avant et après la consonne de liaison.

##### **9.4.1. Les phonèmes se trouvant juste avant la consonne de liaison**

La liaison a moins de raisons d'être réalisée lorsque le mot1 se termine par une consonne prononcée que si le phonème précédant la liaison est une voyelle prononcée. Il y a encore moins de chances que la liaison soit réalisée si le mot1 se termine par deux consonnes prononcées, car cela demande des efforts supplémentaires d'articulation.

Ainsi nous pouvons attribuer les indices suivants :

<b>Terminaison mot1</b>	<b>exemple</b>	<b>indice</b>
Voyelle	Grands enfants	3
Une consonne	Belles histoires	2
Deux consonnes	Tristes histoires	1

**Tableau 3.15. Rôle du contexte phonétique précédant la liaison**

Les phonèmes présents juste avant la consonne de liaison jouent un rôle dans la réalisation de liaisons, et nous pensons que les séquences ayant l'indice 3 ont plus de chances d'avoir une liaison réalisée que celles ayant l'indice 1 (conformément à Delattre, 1955). Toutefois nous nous pencherons dans ce travail davantage sur les questions de fréquence et nous n'aborderons pas cette question phonétique.

#### **9.4.2. Le nombre de liaisons et de syllabes avant et après la consonne de liaison**

L'une des raisons pour lesquelles on ne fait pas ou on fait moins de liaisons aux frontières d'un groupe syntaxique, c'est parce que, comme il y a déjà souvent une ou plusieurs liaisons à l'intérieur d'un groupe syntaxique, cela aurait comme effet qu'il y aurait trop de liaisons à la suite, donc trop de syllabes à la suite à prononcer, ce qui serait trop difficile à produire pour plusieurs raisons.

D'une part, il y a besoin de faire une pause de souffle, généralement toutes les 8-9 syllabes au maximum. Certes, il est possible de prononcer un plus grand nombre de syllabes sans pauses, mais cela demande de les prononcer vite ou même très vite. En parole spontanée en particulier, le temps nécessaire au traitement des informations rendrait cette tâche très difficile.

D'autre part, on fait généralement la liaison entre deux mots au taux de cooccurrence élevé, ou dont la cohésion est (la plus) forte. Ainsi, le taux de cooccurrence est plus fort entre un article et un nom, qu'entre un nom et un adjectif, ou même entre un sujet et un verbe (un groupe nominal et un groupe verbal). Par contre, si le sujet est plus fréquent, comme c'est le cas avec les pronoms, on fait la liaison.

Les liaisons se font le plus souvent entre le mot1 et le mot2, dans ce qu'on pourrait appeler des binômes, et non entre le mot1, le mot2 et le mot3 par exemple (entre des trinômes). Les liaisons vont donc être réalisées entre la séquence (comprenant deux items) au taux de cooccurrence le plus fort.

Cela demande donc de faire des choix, selon des priorités et une hiérarchie, même s'il est théoriquement possible de faire plusieurs liaisons et enchaînements à la suite. Adda-Decker, Boula de Mareüil et Lamel (1999 : 2240) remarquent en effet que les francophones évitent de faire plusieurs liaisons à la suite : « While not prohibited, successive liaisons are generally avoided by speakers ».

Il est théoriquement possible de faire une liaison entre deux groupes syntaxiques, comme par exemple entre un groupe nominal et un complément d'objet (*il met\_un pantalon*) ou entre un groupe nominal et un complément circonstanciel (*ils arrivent\_en avance*), bien que cette liaison soit plus rarement réalisée qu'à l'intérieur d'un groupe syntaxique.

Toutefois, on voit dans le deuxième exemple que, puisque la liaison est obligatoire dans *ils\_arrivent*, et que la liaison est obligatoire dans *en\_avance*, faire trois liaisons à la suite demande quelques efforts parce que les liaisons sont généralement réalisées dans des binômes. Plus le nombre de syllabe est élevé, plus il faut faire des efforts pour ne pas avoir de pause de souffle, et surtout pour réaliser la liaison (ou l'enchaînement) dans autre chose qu'un binôme.

Il est donc moins coûteux, en termes d'efforts, de faire une pause là où c'est possible, c'est-à-dire entre les groupes syntaxiques, et ne pas faire la liaison entre *arrivent* et *en*. Mais si la phrase est courte, que chaque groupe syntaxique est court et qu'il n'y a qu'une seule liaison qui se trouve être entre les deux groupes syntaxiques, la liaison est encore possible.

La réalisation d'une liaison ou d'un enchaînement dépend donc en grande partie de ce qui se trouve avant et après le contexte liant. Par exemple, si l'on prend la même séquence *hôpital où* mais dans deux phrases différentes (dont la seconde est dans le texte de notre étude), on voit qu'il est plus probable que l'enchaînement se fasse dans la première phrase que dans la deuxième phrase :

1. C'est l'hôpital \_où je travaille.
2. Cet\_homme\_ a parlé d'un\_ hôpital \_où il \_y avait de vieux\_ habits\_ et de vieilles\_ éponges.

Dans la première phrase il n'y a aucune séquence liante juste avant la séquence en question (*tal où*), et aucune après. Si un locuteur fait l'enchaînement, il prononcera 8 syllabes à la suite, ce qui est en général le maximum de syllabes prononcées avant une pause de

souffle, mais ce qui est encore possible.

Dans la deuxième phrase il y a deux enchaînements et une liaison juste avant la séquence en question, et un enchaînement et deux liaisons après la séquence. Or, si un locuteur fait les deux liaisons et l'enchaînement dans la première partie de la phrase, il prononcera 9 syllabes à la suite, ce qui est en général le maximum de syllabes prononcées avant une pause de souffle. Donc, le locuteur devrait faire une pause de souffle après *hôpital* d'autant que l'enchaînement entre *il* et *y* est pratiquement obligatoire, ce qui empêcherait de faire une pause plus tard dans la phrase.

Nous venons de voir que s'il y a plusieurs liaisons et enchaînements les uns à la suite des autres, dans une même phrase (ce qui est plutôt rare en situation authentique), le locuteur doit faire une pause de souffle et doit donc choisir où faire la pause. Il devrait d'abord réaliser les liaisons obligatoires, c'est-à-dire celles qui comprennent le plus grand nombre de facteurs favorisants. Cela signifie que si une liaison facultative se trouve dans une phrase qui comprend trop de liaisons (obligatoires) et enchaînements, le locuteur devra faire des choix et devrait plutôt s'arrêter à la fin d'un groupe syntaxique, faire une pause de souffle et faire une liaison ou un enchaînement à l'intérieur du prochain groupe syntaxique.

Dans cette hypothèse des priorités et du contexte, rentre également en jeu, pour les mêmes raisons mentionnées ci-dessus, la longueur des mots. En effet, la longueur des mots participe, au final, au nombre global de syllabes, tout comme le nombre de mots. S'il y a trop de syllabes à la suite, que ce soit parce qu'il y a plusieurs liaisons et enchaînements à la suite, parce que les mots sont longs, ou parce que le groupe nominal est formé de plusieurs adjectifs, cela revient à rallonger le nombre de syllabes et donc à demander des efforts supplémentaires de la part du locuteur, ce qui limiterait les possibilités de liaisons.

Ainsi, lorsqu'on considère le contexte phonétique et prosodique, il est pertinent de prendre en compte, non seulement la séquence elle-même, c'est-à-dire le mot1 et le mot2,



mais aussi les séquences précédant et suivant le contexte liant :

1. les phonèmes présents avant la consonne de liaison
2. les liaisons et enchaînements effectués avant et après la consonne de liaison
3. le nombre de syllabes avant la liaison (et potentiellement après).

Pour cette étude nous n'analyserons pas ces facteurs car nous nous concentrerons sur les effets de fréquence. Nous les mentionnons toutefois pour expliquer que d'autres facteurs contribuent à la liaison, et que les différents facteurs de fréquence ne sont pas les seuls.

## 9.5. La complexité des structures grapho-phonémiques

D'après l'hypothèse de la concurrence et de la hiérarchie, la complexité des structures joue un rôle sur l'acquisition et la production de ces structures. Si une séquence est complexe (les lettres écrites peuvent avoir plusieurs phonèmes correspondants), mais englobe une grande catégorie de mots, cela en favorisera l'apprentissage. Par exemple, puisque *s* se prononce toujours /z/ en liaisons, la structure est peu complexe. Par contre, *f* peut se prononcer /f/ (*neuf enfants*) ou /v/ (*neuf heures*), c'est donc une structure plus complexe.

Si une structure ne comprend pas beaucoup d'items, mais la structure est assez simple, même un nombre de mots restreint dans cette catégorie suffit à l'apprentissage de cette catégorie (la prononciation du *t* de *vingt* dans les chiffres 21-29).

Si une structure est complexe, peu fréquente, et comprend un assez grand nombre d'items cela compromettrait l'acquisition des règles de liaisons et de non-liaisons (les *h* aspirés).

Si une règle phonétique présente dans des structures plus rares (non-liaison avec *h* aspirés) va à l'encontre de règles phonétiques qui s'appliquent à des structures plus fréquentes (liaison dans le contexte [dét. + nom]), on peut considérer qu'elle est complexe.

Nous pouvons donc classer les contextes suivants par ordre croissant d'après la complexité de leurs structures grapho-phonémiques :

1. Structures phonétiques ne présentant aucune difficulté particulière : *Aux invités, les affaires, des étudiants, tes assistants...*

2. Items à prononciation variable et listes restreintes avec possibilité de confusion d'une règle :

- x de liaison qui se prononce z : six heures (dans d'autres cas, se prononce /s/ : *j'en ai six* ou ne se prononcent pas du tout comme dans *six livres*)

- mots commençant par un h muet : possibilité de confusion avec les h aspirés

- d de liaison qui se prononce t : *grand enfant, second étage*

- t de liaison qui se prononce avec une consonne : *vingt-cinq*

3. Structures phonétiques qui regroupent un nombre restreint d'items et pour lesquelles aucune autre règle n'est similaire :

- liaison avec un mot commençant par une semi-consonne : *oie, oiseau, yeux*, ce qui va à l'encontre de la règle selon laquelle il n'y a pas de liaison avec les semi-consonnes (*des yaourts*) (les exceptions des exceptions).

4. Structures phonétiques dont les items sont peu nombreux et vont à l'encontre des règles phonétiques générales :

- *Neuf (v) heures, neuf (v) ans* par rapport à *neuf (f) enfants*.

- Les h aspirés au pluriel les mots au pluriel sont moins utilisés que les mots au singulier (*le héros vs les héros*) ; la règle générale est de faire la liaison, y compris avec les autres h (muets).

- Les semi-consonnes            groupe peu nombreux, pour la plupart d'origine étrangère, mots rares ; la règle s'oppose à la règle générale qui consiste à faire des liaisons.
- Expression figée pour laquelle, en dehors de cette expression, la liaison est facultative ou interdite dans le même contexte syntaxique : *accent aigu* (liaison obligatoire) ; sinon, dans le même contexte syntaxique et phonétique, on ne fait pas la liaison : *un accent étranger* (liaison interdite).
- Liaison interdite avec *et* (liaisons possibles avec d'autres conjonctions: *mais, quand...*)

Il devrait y avoir plus de liaisons produites avec les structures 1 et 2 qu'avec les structures 3 et 4.

### **Exemples de structures phonétiques complexes :**

Certaines séquences présentent des difficultés particulières qui en compliquent le traitement.

La première différence entre l'enchaînement et la liaison, c'est que la prononciation de la consonne de liaison (latente) peut changer, alors que la consonne (fixe) de l'enchaînement est toujours prononcée telle qu'elle est écrite. Par exemple dans *grand enfant* le d de liaison se prononce /t/ alors que dans *grande enfant* le d de liaison se prononce toujours /d/.

Une deuxième différence entre l'enchaînement et la liaison, c'est la possibilité de liaison après une pause. Si l'on fait une pause, en général, on ne fait pas l'enchaînement et on prononce les deux mots séparément. Par exemple, s'il y a une pause après *grande*, on ne sépare pas la consonne d'enchaînement et on ne fait pas l'enchaînement avec /d/. Par contre, il est possible de faire une pause après *les* et de faire la liaison enchaînées avec *z-enfants*.

Une troisième différence entre l'enchaînement et la liaison, c'est que le mot1 a une consonne latente qui autrement qu'en liaison n'est jamais prononcée.

Mais dans *neuf*, on a une caractéristique de l'enchaînement (on ne peut jamais prononcer *neuf* sans le /f/), et deux caractéristiques de la liaison (on peut prononcer f /v/ dans *neuf heures*, *neuf ans*. et on peut faire la pause après la liaison). En fait, c'est un cas très particulier d'enchaînement qui se comporte comme une liaison. Puisqu'il a plus de caractéristiques de la liaison que de l'enchaînement (bien qu'il soit théoriquement un enchaînement) nous l'avons placé dans la catégorie des liaisons. Qu'on l'appelle liaison ou enchaînement, ce qui nous intéressait, c'était surtout d'étudier comment les francophones et les anglophones traitent ce cas particulier.

Une autre difficulté avec *neuf*, ainsi qu'avec *quatre*, *cinq*, *sept*, *huit*, c'est qu'il se termine par une autre lettre que la plupart des adjectifs numéraux (*deux*, *trois*, *six*, *dix*...) dont le pluriel est aussi morphologiquement marqué par /z/. Ainsi, les séquences *neuf heures* et *cent ans*, sont moins « productives » que *deux amis*, *dix ans*, *six heures*, *trois arbres*, *trois ans*, et en plus, la fréquence diminue quand les nombres sont plus grands.

Un autre exemple de structures complexes est les cas de liaisons avec les semi-consonnes. Tout comme avec les h aspirés, ces liaisons vont à l'encontre des règles générales selon lesquelles on fait une liaison avec les déterminants. De plus, à l'intérieur de cette classe restreinte des semi-consonnes, il y a les exceptions *yeux*, *oie* et *oiseau* avec lesquels on fait une liaison. Ces exceptions confirment la règle.

Les mots *yeux* et *oiseau* appartiennent à la catégorie K1, et sont donc très fréquents, peut-être parmi les 100 premiers mots qu'un enfant apprend. Il ne reste qu'à apprendre le mot *oie* (de catégorie K2, donc un peu moins fréquent) qui ressemble d'ailleurs au mot *oiseau* et qui est un *oiseau*. Mais, comme on fait une liaison avec ces items fréquents commençant par une semi-consonne, cela complique la règle, car cette exception consistant à faire les liaisons avec les semi-consonnes donne des exemples qui confirment la règle générale selon laquelle on fait des liaisons avec les déterminants, alors que pour tous les autres items de cette classe

on ne fait pas de liaison.

On pourrait aussi dire que ces mots sont les exceptions des exceptions. Cette catégorie est donc complexe. Elle est toutefois moins complexe que la catégorie des h aspirés, parce que le nombre de mots commençant par une semi-consonne avec lesquels on fait une liaison est très restreint (*yeux, oie et oiseau*), rendant la règle plus générale et plus facile à comprendre et à maîtriser que les règles de liaisons (ou plutôt de non-liaisons) avec les h aspirés.

Autrement dit, alors que pour les règles générales un plus grand nombre d'items rend une structure plus productive, pour les exceptions, c'est le contraire : une catégorie plus restreinte permet de mieux mémoriser ces exceptions. On pourrait presque parler de « productivité des exceptions ». Ainsi la catégorie des exceptions des semi-consonnes (*yeux, oie et oiseau*) devrait être acquise par les francophones parce qu'elle est plus productive que celle des h aspirés (qui contiennent des centaines d'items) tandis que s'il y a des erreurs avec les exceptions, on devrait les trouver dans des catégories contenant plus d'items telles que celle des h aspirés.

En ce qui concerne les expressions figées, celles qui ont le plus de chances d'être retenues sont, comme nous en avons déjà discuté, celles qui sont le plus souvent employées, soit parce l'item lui-même est très employé (*États-Unis, comment allez-vous, vis-à-vis...*), soit parce que ces expressions sont formées à partir des mêmes mots que d'autres expressions partageant un même contexte de liaison (*tout à coup, tout à l'heure, tout un chacun...*), profitant ainsi de l'effet d'analogie.

Une expression telle que *monts et merveilles* pourrait profiter de l'analogie de *ponts et chaussées*, mais comme chacune des expressions est relativement rare, et la productivité de ces deux expressions figées réduite, le taux de réalisation de la liaison devrait être bas.

En conclusion, si l'on prend en considération les différents facteurs phonétiques, nous voyons que le facteur phonétique comprend lui-même des sous-parties. Certaines consonnes

de liaisons favorisent la liaison, à moins que les liaisons ne se trouvent dans des situations complexes telles qu'avec les expressions figées. Enfin, le contexte phonétique est lui-même souvent lié à d'autres facteurs tels que le facteur morphosyntaxique ou prosodique.

## 10. Hypothèses

Ågren (1973) a déterminé les facteurs qui participent à la production de liaisons facultatives dans quarante heures de français parlé: la cohésion, le niveau de langue, la pause hésitation, le coup de glotte, la fréquence d'emploi, la longueur des termes à lier et le facteur phonétique. D'après Ågren (1973), la fréquence d'emploi est l'une des composantes de la cohésion ; et selon lui (1973 : 13) : « L'indice le plus important de cohésion est [...] fourni par la fréquence des L ». De plus, un autre facteur est ce qu'Ågren (1973) nomme « le cas terminal » :

En calculant ces cas terminaux, qui au fond sont des NL causées par l'interdiction de lier à partir d'un mot terminant en GR, je me suis procuré un critère supplémentaire pour évaluer la cohésion. Si une catégorie de séquences en contient peu ou n'en compte pas un seul, c'est que la cohésion est forte. Aussi cette interdépendance se manifeste-t-elle par la fréquence des L et le nombre des cas terminaux qui ne se contredisent jamais l'un l'autre. (Ågren, 1973 : 17)

Ågren explique plus loin (p. 28-29) que « La présence de ce facteur (la fréquence d'emploi) est cependant difficile à démontrer ». En effet, à l'époque d'Ågren et de Delattre, les outils informatiques d'aujourd'hui n'étaient pas encore accessibles aux linguistes. Les banques de données sur le français oral telles que *PFC* ou *Lexique* et les banques de données sur le français écrit telles que *Frantext* ou *Vocabprofil* ne sont accessibles en ligne que depuis quelques années. Fougeron et al. (2001b) mentionnent eux aussi le rôle de la fréquence :

Pour qu'il y ait liaison, le lieur et le lié doivent former une paire de catégories syntaxiques valides. Pour certaines d'entre elles, la liaison est considérée comme obligatoire [...]. Pour d'autre, elle est facultative [...], et c'est alors la fréquence de la cooccurrence de la paire qui, entre autre, influencera la réalisation de la liaison (p. ex. « relations amoureuses »). (Fougeron et al., 2001b : 3)

D'après le MBU, le cerveau humain analyse les informations d'après un système complexe de probabilités basées sur la fréquence de plusieurs facteurs dont la fréquence des items et la fréquence des structures. Aussi, nous souhaitons analyser ici certains facteurs, tous en rapport avec la fréquence, en tentant de reproduire quelques données de ce système de traitement des informations. Nous nous concentrerons au prochain chapitre sur les facteurs de fréquence et analyserons le rôle de la fréquence dans les productions des francophones et des anglophones.

Notre hypothèse principale est la suivante : si au moins deux facteurs de fréquence sont élevés (par exemple, la fréquence de la structure grammaticale et la fréquence de cooccurrence des deux mots), la production de liaisons devrait l'être également. Si tous les facteurs sont bas, le taux de production devrait l'être également, à moins qu'un élément perturbateur (tel qu'une possibilité d'homophonie) l'en empêche. Nous analyserons chaque facteur séparément avant de les analyser conjointement. Nous calculerons les tendances générales avec des pourcentages, avant de raffiner l'analyse avec des calculs statistiques.

D'après les études dont nous avons discuté dans les chapitres précédents, nous nous attendons à ce que les liaisons obligatoires soient réalisées par les francophones majoritaires à plus de 95%, et les liaisons facultatives entre 0 et 40%, selon les contextes syntaxiques, phonétiques et lexicaux. Nous nous attendons aussi à ce qu'il y ait des liaisons interdites réalisées, par analogie avec les règles générales, en particulier par les groupes minoritaires. Nous nous attendons également à ce que les enchaînements soient produits en plus grande quantité chez les groupes majoritaires que chez les groupes minoritaires.

Selon le MBU, plus l'input a été riche en début d'acquisition et dans la vie du locuteur, plus le locuteur aura « enregistré » de séquences, plus riche sera sa banque de données, et plus le locuteur pourra traiter les informations selon un modèle qui se conforme aux normes orthoépiques. C'est-à-dire que plus un sujet aura entendu de liaisons, plus il sera à

même de les produire et vice-versa : moins il aura entendu certaines liaisons, moins il les produira.

On a vu avec l'acquisition des liaisons des enfants L1 (Dugua, 2005, 2006 ; Chevrot, Chabanal et Dugua, 2007) que les liaisons les plus fréquentes sont les plus vite acquises et que les liaisons obligatoires sont acquises à 6 ans, tandis que les liaisons facultatives ne sont pas complètement acquises même à 11-12 ans, car elles sont moins fréquentes (Pallaud et Savelli, 2001) et leur prononciation varie. Les séquences les plus fréquentes sont le plus vite acquises et celles dans lesquelles il y a le moins d'erreurs.

Lorsque les sujets ont moins accès à la langue, ils produisent moins de liaisons parce qu'ils en ont sans doute moins entendu (les Africains du projet PFC par exemple). On sait que, si les locuteurs n'ont pas (assez) entendu de séquences, ils ne les ont pas enregistrées dans leur lexique et ne font donc pas les liaisons (Chabanal et Embarki, 2002). S'ils les ont entendues un grand nombre de fois ils les ont assimilées et produisent les liaisons. On peut donc en conclure qu'il pourrait y avoir un seuil de fréquence nécessaire à l'assimilation de structures et séquences, seuil au-dessus duquel la fréquence n'aurait plus d'importance puisque les séquences seraient enregistrées.

Ainsi, qu'un locuteur ait entendu des séquences 2000 fois ou 10000 fois ne changerait pas le traitement des données, puisqu'à partir d'un nombre de séquences, la structure serait « enregistrée » ou « enracinée ». De plus, la fréquence joue un rôle plus important pour les structures plus complexes ou pour les exceptions, comme par exemple les expressions figées et les liaisons interdites avec les h aspirés.

S'il est difficile de déterminer avec exactitude combien de fois un locuteur a entendu une séquence, on peut tout de même considérer plusieurs facteurs qui vont participer aux probabilités d'en avoir entendu certaines. Parmi ces facteurs, on compte, d'une part, les facteurs linguistiques propres aux séquences, donc les divers types de fréquence : la fréquence



du Mot1, la fréquence du Mot2, la fréquence de cooccurrence des Mot1 et Mot2 et la fréquence de la structure. On compte d'autre part, les facteurs sociolinguistiques qui définissent le capital linguistique des locuteurs : l'accès du locuteur à la langue, en particulier en début d'acquisition.

Ainsi, nous supposons que les francophones majoritaires et minoritaires produiront de façon systématique les liaisons regroupant le plus grand nombre de facteurs de fréquence et que seuls les francophones majoritaires ayant été exposés à un input riche et varié réaliseront les liaisons obligatoires moins fréquentes.

Nos hypothèses sont les suivantes :

1. Les liaisons les plus fréquentes devraient être produites par tous les francophones (90% et + par tous les groupes) et dans une très grande proportion par les apprenants L2.
2. Les liaisons moins fréquentes (les adjectifs + noms, et certaines expressions figées) devraient être produites en moins grand nombre que les liaisons plus fréquentes, mais quand même en grand nombre (70-90%) par les francophones majoritaires, en plus faible proportion par les francophones minoritaires (50-70%), et en encore plus faible proportion par les apprenants (moins de 50%).
3. La fréquence des items devrait jouer un plus grand rôle pour les liaisons moins fréquentes et les différences devraient être les plus visibles pour les groupes minoritaires avec les liaisons moins fréquentes (pour les liaisons plus fréquentes, passé un seuil de fréquence, il n'y a plus de différence, une fois les séquences enregistrées).
4. La fréquence du mot1, du mot2 et la fréquence de cooccurrence devraient jouer un rôle majeur dans la production de liaisons, ainsi que la consonne de liaison.

5. La productivité des structures devrait jouer un rôle dans la production de liaisons. Si les règles d'une structure vont à l'encontre de structures plus productives et que, de surcroît, ces items sont moins fréquents, alors l'effet cumulé du taux réduit de fréquence, du nombre réduit d'items et de la règle allant contre les règles plus générales limiterait la production de liaison. Les liaisons dans les séquences qui regroupent le plus grand nombre de facteurs « favorisant » la liaison devrait être réalisée par tous les francophones en plus grand nombre que les contextes contenant moins de facteurs.
6. Les mêmes tendances devraient être observées pour les liaisons facultatives et les enchaînements, et il devrait y avoir un continuum entre les francophones majoritaires, bilingues, minoritaires et apprenants L2.
7. Les apprenants du français langue seconde ou langue étrangère (il y a quelques sujets non canadiens) ne devraient pas maîtriser le système des liaisons comme les francophones majoritaires. Ils ne traiteraient pas les informations selon les mêmes critères à la fois syntaxiques, phonétiques et lexicaux que les francophones.
8. Nous supposons que les apprenants n'ont, pour la plupart, pas développé de catégories abstraites et ne sont pas encore capables de produire des séquences jamais entendues ou peu entendues. À ce stade de leur développement, ils peuvent reproduire les séquences les plus souvent entendues, qu'elles soient obligatoires ou facultatives. Les liaisons ne sont systématiquement réalisées (à plus de 95%) dans aucune catégorie syntaxique même dans la catégorie syntaxique la plus fréquente [dét. + nom]. Les apprenants traitent les items et les séquences au coup par coup, mais ne sont pas capables d'appliquer les informations à des catégories générales, de façon systématique.

9. Nous supposons qu'ils produisent un nombre beaucoup plus bas d'enchaînements, de liaisons obligatoires, de liaisons facultatives que les francophones, et produisent plus de liaisons interdites. Ils produisent aussi beaucoup plus de liaisons non-enchaînées, et font plus d'erreurs que les francophones.
10. Nous nous attendons à ce que les liaisons obligatoires soient réalisées en plus faibles proportions que dans les études de Mastro Monaco (2000), Thomas (2002) et Howard (2005) (dans ces études, les productions étaient supérieures à 90% de réalisations de liaisons obligatoires).
11. Nous pensons que la langue maternelle joue un rôle dans la production de liaisons. Si des sujets ont déjà appris plusieurs langues, ils pourraient être plus sensibles à certains aspects de la langue, notamment les aspects phonétiques. Par exemple, des sujets parlant déjà leur langue maternelle, plus l'anglais L2 et le français L3, pourraient avoir un taux de réalisation plus élevé que les personnes qui n'ont que le français comme langue étrangère ou langue seconde, notamment parce qu'ils ont un plus grand répertoire phonétique dans lequel puiser.
12. Le groupe qui a entendu le plus de séquences (au moment de l'enregistrement) devrait produire un plus grand nombre de liaisons que les autres groupes.
13. Tous les groupes devraient faire des progrès après la leçon sur les liaisons parce que les sujets se sont familiarisés avec le texte et ont donc pratiqué la lecture de ces séquences.
14. Certains groupes pourraient faire plus de progrès que d'autres, car les groupes ont reçu plus de répétitions que d'autres, certains des explications plus longues que d'autres, etc.

15. Les expériences linguistiques des sujets pourraient influencer sur la production de liaisons (que nous vérifierons avec les réponses au questionnaire).
16. Les informations phonétiques qu'utilisent les anglophones pour le traitement des liaisons sont différentes de celles qu'utilisent les francophones, parce que les anglophones n'ont pas les mêmes références phonétiques et subissent l'influence de leur L1. Ils n'ont pas eu assez d'input en L2 pour « perfectionner » leurs connaissances phonétiques en général, et le système de syllabation et des liaisons en particulier.
17. Le type d'enseignement et le type d'école jouerait un rôle dans la production de liaisons.
18. Le nombre d'années passées à étudier le français ne garantirait pas des productions supérieures. En effet, si les apprenants n'ont eu aucun cours de prononciation, n'ont pas assez entendu de séquences liantes, et s'ils se concentraient sur le message et non sur la forme, ils pourraient ne pas faire plus de liaisons que ceux qui ont étudié le français moins longtemps. Si leur prononciation s'est très tôt fossilisée, après un certain temps, le nombre d'années d'études n'améliorerait plus la prononciation (dont les liaisons).
19. Nous pensons que si les étudiants participent à de nombreuses activités en français telles qu'écouter le français avec des chansons, la télévision, la radio, etc., ils auront une fréquence d'écoute plus grande, un input plus riche et feront plus de liaisons que ceux qui sont moins en contact avec la langue. De plus, participer à ces activités exprime la motivation des étudiants à s'améliorer et à faire attention à la forme.
20. Les brefs séjours à l'étranger n'auraient aucun effet sur les productions de liaisons. Le temps d'exposition à la langue cible était trop court et était insuffisant pour améliorer la banque de données.

21. Enfin, la proportion devrait être inversement proportionnelle pour les liaisons interdites. Les apprenants devraient en faire le plus (ils maîtrisent assez bien les cas les plus fréquents, mais probablement moins les exceptions tels que les h aspirés), suivis des francophones minoritaires, et enfin les francophones majoritaires devraient en faire en plus petite quantité que les autres groupes, probablement moins de 5%.
22. D'après l'attention portée sur la forme et la théorie de la fréquence, les étudiants devraient tous faire des progrès, mais il devrait y avoir aussi des différences personnelles dépendant des aptitudes personnelles de chacun, des méthodes pédagogiques utilisées au moment du test, mais aussi des différences selon les items du test. Il devrait y avoir plus de progrès pour les liaisons des mots les plus fréquents. Les mots les plus rares devraient être ceux pour lesquels les progrès seraient les plus lents.
23. Il est aussi probable que les étudiants qui font le plus de progrès, soient ceux qui étaient les plus faibles au départ, comme c'était le cas dans l'étude de Cenoz et Lecumberri (1999) portant sur la discrimination des voyelles en anglais.
24. Cette étude devrait être en accord avec les études précédentes sur le MBU. La fréquence serait un facteur majeur dans la production des francophones et des anglophones.

Ainsi, selon le MBU, la fréquence d'écoute et de production de liaisons des locuteurs, en particulier (mais pas uniquement) en début d'acquisition, influence l'acquisition des liaisons. Plus l'input a été riche, plus la banque de données est riche et variée, plus le système d'analyse des données peut être performant. Nous avons vu cela pour l'acquisition des liaisons chez les enfants (Tomasello, 2000a, 2000b; Ellis, 2002).

En outre, les personnes étant en contact permanent avec des francophones en milieu majoritaire, ou l'ayant été jusqu'à l'âge adulte, dans toutes les situations de la vie courante (à

l'école, à l'université, à la maison, avec les amis, les voisins, pour les loisirs, etc.) auront entendu plus de situations « liantes » que des personnes s'étant trouvées en contact avec des francophones dans des situations restreintes (français entendu seulement à la maison ou seulement à l'école). Selon le MBU, le nombre d'heures de français entendu en début d'apprentissage joue un rôle sur l'acquisition. Ici, l'input reçu influencerait sur le nombre de liaisons entendues et donc produites.

## Conclusion

Les facteurs précédemment mentionnés sont liés les uns aux autres. Si des items sont très fréquents, ils vont probablement créer une classe. Si cette classe contient de nombreux items, elle est productive. Dans ce cas, la structure syntaxique ou phonétique est productive. Si certains contextes phonétiques ou syntaxiques handicapent la liaison, la liaison est amenée à disparaître dans ces contextes.

Les liaisons sont faites dans les séquences contenant les mots les plus fréquents. C'est la raison pour laquelle il est difficile de considérer uniquement le facteur syntaxique ou phonétique. Par exemple dans la catégorie [adj. + nom], certains adjectifs sont très fréquents (*vieux, bon, etc.*) alors que d'autres sont beaucoup plus rares. On trouve aussi des différences de fréquence avec *Lexique 3* quant au genre, au nombre, et aux personnes.

Pour mieux comprendre le système, il est utile de créer un système probabiliste qui reproduise, autant que possible, le système de traitement des informations du cerveau. Plus les facteurs favorisant la liaison sont présents, plus les probabilités sont élevées. Plus les probabilités sont élevées, plus la liaison est susceptible d'être produite.

C'est ce que nous entendons tester à l'aide de *GoldVarb* dans le prochain chapitre.

D'après ce système de probabilités, nous émettons l'hypothèse que les enchaînements, liaisons obligatoires, liaisons facultatives dont les facteurs favorisants sont les plus élevés sont ceux qui ont le plus de probabilités d'être produites par tous les francophones. C'est ce que nous allons tester dans les productions de francophones.

Les liaisons qui réunissent moins de facteurs ou dont les facteurs sont peu élevés ne devraient être réalisées par personne, et les réalisations des liaisons facultatives se trouvant entre ces deux extrêmes devraient dépendre du capital linguistique des locuteurs, et inversement pour les liaisons interdites.

Mallet (2008) a analysé la production de liaisons dans la lecture d'un texte, la parole semi-guidée et la parole spontanée dans le corpus PFC. L'auteure conclut sa thèse en mentionnant qu'il faut analyser plusieurs facteurs en même temps, notamment le facteur de longueur, les facteurs externes et internes (âge du locuteur...) pour mieux comprendre les liaisons. C'est précisément ce que nous ferons ici. Nous analyserons plusieurs facteurs en même temps.

## **CHAPITRE IV**

### **RÉSULTATS ET ANALYSES (POURCENTAGES)**



## 1. Les productions de liaisons et enchaînements des francophones

Le but de notre étude est d'évaluer les connaissances et compétences des anglophones. Pour se faire, nous avons besoin d'un groupe témoin. Afin de comparer les productions de liaisons d'anglophones à celles de francophones, il est utile de comparer la même chose, c'est-à-dire de comparer des francophones lisant le même texte que les anglophones.

Le but principal de cette étude est de tester les liaisons obligatoires les plus fréquentes à l'intérieur du groupe nominal. Les données sur les liaisons facultatives et interdites sont surtout présentes afin de fournir des renseignements complémentaires afin d'analyser certaines tendances remarquées avec les liaisons obligatoires. Les enchaînements sont notamment présents dans le texte afin de déterminer les difficultés relatives à la syllabation, en particulier pour les anglophones.

Nous allons ici évaluer les productions de liaisons et enchaînements des francophones majoritaires et minoritaires et étudier le rôle des facteurs linguistiques et en particulier de la fréquence sur la production de liaison en utilisant le logiciel *GoldVarb*. Toutefois, vu le nombre restreint de sujets dans chaque catégorie, cela nous permettra d'établir des tendances générales, mais pas de tirer des conclusions sur chaque groupe de francophone. Ceci n'est pas le but de cette thèse, car celle-ci porte principalement sur les sujets anglophones. Les sujets francophones sont présents dans l'étude afin de mieux comprendre et interpréter les résultats des apprenants anglophones et le fait d'avoir différents groupes francophones permet d'éviter d'avoir un modèle dialectal unique.

## 1.1. Les études sur les productions de francophones

Il est difficile de comparer les chiffres des études précédentes à ceux de notre étude car elles ne comparent pas exactement les mêmes choses. Aussi, nous comparerons les seuls faits comparables quand c'est possible : la liaison en contexte [dét. + nom], et la liaison en contexte [adj. + nom]. Même dans ces contextes, les liaisons n'ont pas été catégorisées et calculées de la même manière, ce qui nous pousse à considérer ces données avec précaution.

Rappelons d'abord quelques études abordées précédemment. Selon Malécot (1975) aucune liaison n'est réalisée à 100% sauf avec la préposition *en*. La liaison est réalisée à 98% avant un nom (entre déterminant et nom), à 99.5% après tous les articles, à 98.6% après un pronom personnel sujet, à 98.8% après une préposition, à 93.2% après un adjectif. Selon Booij et de Jong (1987), les liaisons sont réalisées à 100% après un déterminant, après un pronom préverbal, entre un verbe et un pronom (*vas-y*, *prenez-en*) et dans certaines expressions figées (*tout à l'heure*).

Les études de Durand, Laks et Lyche (2008) et de Mallet (2008) sur les liaisons du projet PFC, ainsi que celles de Boula de Mareüil, Adda-Decker et Gendner (2003) et Fougeron et Delais-Roussarie (2004), considèrent si les liaisons sont réalisées ou non, ainsi que les conditions de réalisation (avec ou sans enchaînement, hésitation, etc.). Par contre, les liaisons ne sont pas divisées par catégories (liaisons obligatoires, facultatives ou interdites). C'est pourquoi les pourcentages de liaisons de ces études sont plus bas que ceux de Malécot (1975), Ågren (1973) ou Booij et de Jong (1987).

Pour Boula de Mareüil, Adda-Decker et Gendner (2003), les liaisons sont réalisées à 95.3% après un déterminant, à 90.5% après une préposition monosyllabique et à 72.5% entre un adjectif et un nom. Pour Durand, Laks et Lyche (2008) et Mallet (2008), les liaisons (obligatoires et facultatives) sont réalisées et enchaînées (résultats retrouvés le 24 octobre

2010 sur le site du projet PFC<sup>41</sup>) dans 59.37% des cas, non-réalisées dans 39.51% des cas, réalisées, mais non-enchaînées, dans 0.78% des cas si l'on considère tous les sujets (L1 et L2). Si l'on considère seulement les francophones de France, on obtient 62.6% de liaisons réalisées et enchaînées, 36.28% de liaisons non-réalisées, et 0.73% de liaisons réalisées, mais non-enchaînées. Le taux de liaisons réalisées est plus élevé en lecture qu'en conversation. Les résultats des francophones canadiens ne sont pas encore disponibles.

L'un des buts du projet PFC est d'analyser la variation (les réalisations de liaisons et de schwas). Parmi les différents types de variation, on compte notamment la variation géographique. Par exemple, en lecture, les liaisons sont réalisées et enchaînées à 66.04% à Lacaune (France Midi-Pyrénées), à 56.82% en Vendée et à 36.25% au Cameroun. Les productions varient presque du simple au double. Notons que pour les groupes africains, le français est une langue seconde, mais c'est leur langue d'éducation et ils la maîtrisent à un niveau très proche des francophones majoritaires.

En lecture, mis à part le pourcentage plus élevé pour Abidjan qui est de 61.82%, tous les autres groupes africains ont un taux de réalisations plus bas que les groupes européens. Le groupe de Bangui compte 46.48 % de liaisons enchaînées et le groupe camerounais 36.25%. Le seul groupe qui ait un taux élevé de liaisons non-enchaînées qui reflète autre chose que de simples hésitations, est le groupe camerounais, avec 13.75% de liaisons non-enchaînées.

Le nombre élevé de liaisons non-enchaînées associé au nombre réduit de liaisons enchaînées, suggère la présence d'un système phonétique différent du français standard et de celui des francophones majoritaires. Voyons maintenant les résultats de notre étude.

---

<sup>41</sup> [http://www.projet-pfc.net/index.php?option=com\\_wrapper&view=wrapper&Itemid=148](http://www.projet-pfc.net/index.php?option=com_wrapper&view=wrapper&Itemid=148)

## **1.2. Résultats de notre étude**

Nous analyserons ici les tendances générales de productions de liaisons et enchaînements des francophones. Puis nous expliquerons certains phénomènes, notamment le rôle de la fréquence dans la production de liaisons. Nous avons, dans ce chapitre, fait des calculs généraux avec des pourcentages. Les hypothèses amorcées avec les pourcentages sont analysées avec des statistiques dans le dernier chapitre.

Rappelons que le but de notre étude était d'abord de tester le rôle des différents types de fréquence dans la production de groupes francophones, et surtout d'étudier les productions des anglophones pour lesquels les francophones sont des groupes témoins. Nous testons néanmoins certains facteurs, pour déterminer ceux qui semblent plus pertinents que d'autres pour notre analyse.

Lorsque nous utilisons le terme « norme », cela s'applique aux liaisons obligatoires et interdites. Quant aux liaisons facultatives, la seule norme dont il s'agit ici, c'est la façon de prononcer la séquence, avec le bon phonème avant, dans et après la liaison, Sinon, nous parlons de réalisation (« correcte », avec la bonne consonne de liaison et l'enchaînement). Une non-réalisation de liaison facultative n'est bien sûr pas considérée comme une erreur.

### **1.2.1. Résultats des productions de francophones**

Dans bon nombre d'études, les facteurs sociaux tels que le sexe, l'âge et le niveau d'étude jouent un rôle dans la production de liaisons (par exemple, Encrevé, 1988). Toutefois, si plusieurs facteurs sont réunis, tels que la fréquence d'écoute de la langue, la fréquence des items et des structures, alors ces facteurs sociaux semblent jouer un rôle moins déterminant que d'autres facteurs, ce qui n'est d'ailleurs pas le sujet de notre étude. En outre, les facteurs sociaux propres aux locuteurs dépendent souvent les uns des autres, si bien qu'il est difficile d'en interpréter les résultats.

### 1.2.1.1. Le sexe des sujets

Le sexe des sujets semble faire une légère différence dans la production de liaisons (tableau 4.1). Les hommes produisent 94.4% de liaisons obligatoires et les femmes 89.7%. Toutefois, étant donné les dimensions limitées de notre groupe témoin de francophones, le sexe des sujets correspond aussi à une situation linguistique (tous les sujets franco-ontariens étaient des femmes), et d'autres facteurs tels que le niveau d'étude, par exemple, tous les hommes sauf 1 ont fait au moins 5 ans d'études universitaires. Il est ainsi difficile de séparer le facteur sexe des autres facteurs et de tirer des conclusions de ce facteur.

	Hommes			Femmes		
	Nb	total	%	Nb	Total	%
LO	385	408	<b>94.4</b>	549	612	<b>89.7</b>

**Tableau 4.1. Productions de liaisons obligatoires par sexe**

### 1.2.1.2. L'âge des sujets

Dans notre étude (tableau 4.2.), on ne peut pas considérer que l'âge est un facteur déterminant dans la production de liaisons, dans la mesure où l'âge est plus relié à la situation linguistique qu'à tout autre facteur. Par exemple, tous les participants francophones de plus de 60 ans étaient des personnes franco-ontariennes en situation d'emploi restreint de la langue. Ainsi, la plus faible proportion des sujets de plus de 60 ans est plus reliée à la situation linguistique qu'à l'âge. C'est la raison pour laquelle, les productions de liaisons ne sont pas corrélées aux autres tranches d'âge.

âge	nb	total	%
18-25	183	204	<b>89.7</b>
26-29	184	204	<b>90.2</b>
30-39	245	255	<b>96.1</b>
40-49	146	153	<b>95.4</b>
50-59	0	0	<b>0</b>
60+	176	204	<b>86.3</b>
total	934	1020	<b>91.6</b>

**Tableau 4.2. Age et production de liaisons obligatoires**

### 1.2.1.3. Le niveau d'étude

Le niveau d'étude est lui aussi lié à d'autres facteurs (tableau 4.3.). Par exemple trois participants franco-ontariens sur quatre ont fait moins de quatre années d'études universitaires, ce qui est moins que la plupart des autres locuteurs. Les personnes ayant obtenu leur doctorat étaient, pour la plupart, des francophones majoritaires. Il est donc difficile d'attribuer les chiffres suivants uniquement à l'âge, puisque l'âge est aussi directement relié au type de contact avec la langue. Moins les personnes ont eu de contact avec la langue (franco-ontariens et certains Africains), moins elles ont fait d'études universitaires (sans que cela en soit une cause ou une conséquence). La plupart des personnes ayant fait plus de 4 années d'université étaient des francophones majoritaires.

Nb années d'études universitaires	nb	total	%
0 :	0	0	<b>0</b>
1-2 :	45	51	<b>88.2</b>
3-4 (BA):	169	204	<b>82.8</b>
5 (Master):	429	459	<b>93.5</b>
6 (DEA) :	93	102	<b>91.2</b>
7+ (doctorat) :	198	204	<b>97.1</b>
total	934	1020	<b>91.6</b>

**Tableau 4.3. Années d'études universitaires et production de liaisons**

Tous ces facteurs sociaux sont difficiles à interpréter parce que les conditions d'apprentissage sont différentes entre les locuteurs. Pour identifier l'effet de l'âge ou du niveau d'étude, il faudrait que les autres conditions soient au moins similaires. Ici, ces facteurs ne semblent pas jouer un rôle majeur.

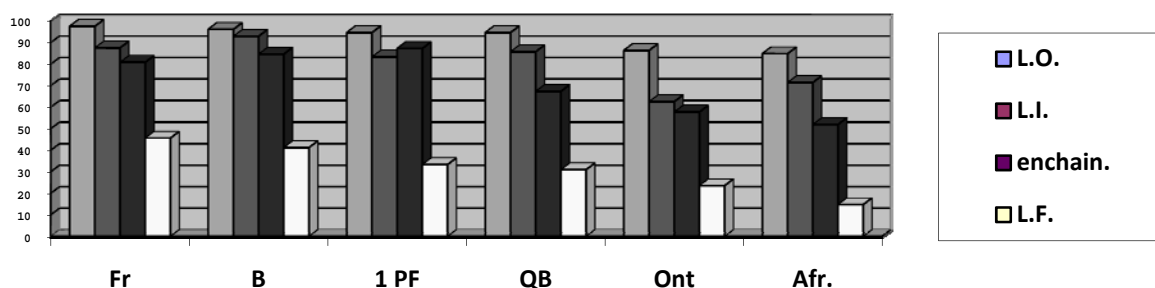
### 1.2.2. Résultats généraux

Les tendances semblent être les mêmes pour tous les groupes de notre étude (graphique 4.1) en ce qui concerne la proportion de liaisons obligatoires, facultatives et interdites, ainsi que la proportion d'enchaînements, les uns par rapport aux autres. C'est le pourcentage de chaque catégorie qui varie légèrement d'un groupe à l'autre.

Les liaisons **obligatoires** sont réalisées en grande majorité chez tous les groupes, dont 95.6% pour les francophones majoritaires et 85.5% pour les francophones minoritaires. Les **enchaînements** sont réalisés dans plus de la moitié des cas par tous les groupes : à 78.2% pour les francophones majoritaires et à 54.8% pour les francophones minoritaires.

Les liaisons **interdites** sont produites dans une faible proportion par tous les groupes. Leur prononciation se conforme à la norme et les liaisons interdites ne sont **pas** produites dans 86.3% des cas pour les francophones majoritaires et dans 67% pour les francophones minoritaires. En d'autres termes, les francophones ont réalisé 13.7% de liaisons interdites et les francophones minoritaires en ont produit 33%. Les francophones minoritaires ont réalisé presque trois fois plus de liaisons interdites que les francophones majoritaires.

Enfin, les **liaisons facultatives** sont réalisées dans une plus faible proportion que les liaisons obligatoires et les enchaînements par tous les groupes : à 37.3% pour les francophones majoritaires et à 19.1% pour les francophones minoritaires.



Graphique 4.1. Production de liaisons et enchaînements tous francophones

Il est difficile, voire impossible, de réaliser à la fois les 51 liaisons obligatoires, les 17 liaisons facultatives et les 13 enchaînements dans notre texte. Cela constituerait un total de 81 cas liants (liaisons + enchaînements) sur 236 mots. Si les sujets faisaient toutes les liaisons obligatoires et facultatives et tous les enchaînements, cela les forcerait à prononcer au moins 18 ou 20 syllabes à la suite dans chaque phrase, ce qui, bien que possible, ne permettrait pas de pauses de souffle et demanderait aux sujets de lire à une rapidité extrême, alors qu'il est précisé qu'ils doivent lire à vitesse « normale ».

Les sujets ont donc dû choisir parmi ce qui leur paraissait plus « naturel » et « prioritaire » en coupant les groupes de mots et en ne réalisant pas certaines liaisons et enchaînements pour leur permettre des pauses de souffle et leur permettre de lire à vitesse normale. Ceci explique que, si une catégorie est plus élevée chez un sujet, cela puisse faire baisser une partie de l'autre catégorie. Par exemple, si un sujet réalise une grande proportion de liaisons obligatoires et d'enchaînements, cela pourrait réduire la proportion des liaisons facultatives réalisées, si celles-ci se trouvent juste après d'autres liaisons effectuées.

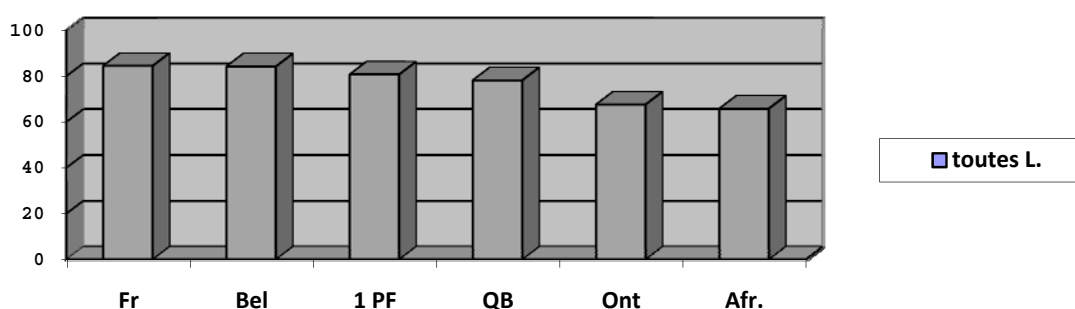
De même, si un sujet réalise un grand nombre de liaisons facultatives, son nombre d'enchaînements pourrait être plus bas. Par exemple, la phrase « les personnes ont écrit des histoires sur ces hôpitaux » comprend 14 syllabes. Il y a une liaison obligatoire entre *des histoires* et *ces hôpitaux*, et *ces hôpitaux* fait partie du même groupe syntaxique (complément d'objet). Pour prononcer 14 syllabes à la suite, il faudrait faire un effort particulier, tel que prendre une grande respiration avant de commencer, sinon les locuteurs ont besoin de faire une pause de souffle.

Ainsi, si les locuteurs font une pause, ils doivent choisir entre faire l'enchaînement dans *les personnes ont*, et faire la liaison facultative dans *ont écrit*. Il est possible que, face à un tel dilemme, les locuteurs choisissent une catégorie plutôt qu'une autre, auquel cas la catégorie choisie (par exemple les liaisons facultatives), serait produite en plus grand nombre



que l'autre cas de sandhi facultatif, l'enchaînement.

C'est la raison pour laquelle il est utile de considérer également la somme de tous les cas de sandhi : de toutes les liaisons et de tous les enchaînements pour déterminer la proportion globale (graphique 4.2.) Ainsi, alors qu'il y a peu de différences entre les groupes majoritaires pour les liaisons obligatoires (97.5, 96.1, 94.1, 94.6 pour les groupes de France, Belgique, de parent francophone, et du Québec), de légères différences se dessinent pour les productions globales avec 84.47%, 84.21% ; 80.7% et 78.15% pour les mêmes groupes.



**Graphique 4.2. Productions de toutes liaisons et enchaînements, tous francophones**

Notons que la façon dont nous avons codé les liaisons interdites est différente des autres études dans le domaine pour les raisons suivantes. Dans les autres études, les auteurs ont compté le nombre de liaisons interdites réalisées. Disons, par exemple 7 liaisons réalisées sur 56 sites possibles, soit 12%.

Si l'on veut ajouter les liaisons réalisées dans toutes les catégories et obtenir ainsi une moyenne, et que l'on calcule par exemple (résultats de notre étude pour les Français)

199	liaisons obligatoires	/ 204
+	31 liaisons facultatives	/ 68
+	7 liaisons interdites sur	/ 56

Cela fait un total de 237/328, ce qui ferait une moyenne de 72.25%.

Toutefois, à notre avis, ce chiffre n'est pas représentatif car, ce qu'on veut vraiment calculer, c'est ce qui correspond aux « normes orthoépiques ». Or, il semble que lorsque les liaisons interdites ne sont pas réalisées, cela fait baisser « la moyenne », alors qu'en réalité, cela devrait faire augmenter la moyenne, puisque les productions se rapprochent de la norme qui consiste à ne pas réaliser les liaisons interdites.

Ainsi, nous avons considéré que dans le même contexte que celui ci-dessus

199	liaisons obligatoires	/ 204
+ 31	liaisons facultatives	/ 68
+ 49	liaisons interdites sur	/ 56

49 n'ont pas été réalisées, donc produites « correctement » d'après la norme des liaisons interdites. Cela fait un total de 279/328, ce qui fait une moyenne de 85.06% de liaisons réalisées selon les normes (au lieu de 72.25%).

Les plus grandes différences entre les groupes majoritaires sont le taux d'enchaînements et de liaisons facultatives (tableau 4.6.). Le groupe québécois a produit 67.3% d'enchaînements, alors que les trois autres groupes majoritaires en ont produit entre 80.8% et 87.2%. Le groupe québécois a produit 30.9% de liaisons facultatives, le groupe bilingue 33.3%, le groupe de France 45.6% et de Belgique 41.2%.

Pour les liaisons interdites, si l'on ajoute les liaisons des trois premiers groupes francophones, on obtient un total de 97 liaisons interdites sur 112, soit 86.6%, par rapport à 85.7% pour le groupe québécois. Il y a peu de différences entre les groupes majoritaires pour les liaisons interdites. Étant donné que les normes orthoépiques s'appliquent principalement aux liaisons obligatoires et interdites, on peut considérer que les productions de tous les groupes majoritaires se rapprochent des normes.

Comme le mentionnent Boula de Mareüil, Adda-Decker et Gendner (2003 : 1331), on ne peut pas dire que les liaisons obligatoires telles que décrites dans les manuels de phonétique sont réalisées à 100%, ni que les liaisons interdites ne sont jamais réalisées.

tous	Fr.			Bel.			Bil.			Queb.			total Franc Maj		
	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
Ench.	42	52	<b>80.8</b>	11	13	<b>84.6</b>	34	39	<b>87.2</b>	35	52	<b>67.3</b>	122	156	78.2
L. O.	199	204	<b>97.5</b>	49	51	<b>96.1</b>	144	153	<b>94.1</b>	193	204	<b>94.6</b>	585	612	95.6
L. F.	31	68	<b>45.6</b>	7	17	<b>41.2</b>	17	51	<b>33.3</b>	21	68	<b>30.9</b>	76	204	37.3
L. I.	49	56	<b>87.5</b>	13	14	<b>92.9</b>	35	42	<b>83.3</b>	48	56	<b>85.7</b>	145	168	86.3
<b>Total</b>	321	380	84.4	80	95	84.2	230	285	80.9	297	380	78.1	928	1140	81.4

**Tableau 4.4. Production de liaisons et enchaînements francophones majoritaires**

Si l'on compare les productions des groupes franco-ontariens et africains, on observe des tendances très similaires (tableau 4.5.). Leur production globale est très proche : 67.63% pour le groupe franco-ontarien et 65.78% pour le groupe africain. Les productions à l'intérieur de chaque catégorie sont également très proches. Les enchaînements sont réalisés à 57.7% pour les Franco-ontariens et à 51.9% pour les Africains. Les liaisons obligatoires sont réalisées à 86.3% pour les Franco-ontariens et à 84.8% pour les Africains. Les liaisons facultatives sont réalisées à 23.5% pour les Franco-ontariens et à 14.7% pour les Africains. Enfin, les liaisons interdites ont été correctement prononcées (donc non-réalisées) à 62.5% pour les Franco-ontariens et à 71.4% pour les Africains.

Rappelons que notre petit groupe « franco-ontarien », de par sa taille réduite, (quatre sujets) et de par son âge (plus âgé que le reste des sujets) et en raison de l'exposition des sujets au français (exposition réduite), n'est pas représentatif de la majorité des Franco-ontariens qui ont aujourd'hui accès à des écoles francophones et à de nombreuses activités en français. Le but de cette étude est d'analyser les productions de personnes ayant eu un contact plus restreint avec la langue. Nous n'en tirons en aucun cas des conclusions sur tous les Franco-ontariens. Les groupes franco-ontariens et africains de notre étude ont eu un contact plus restreint avec la langue française que les groupes majoritaires et cette différence d'input se reflète dans les productions proportionnellement restreintes de liaisons réalisées.

	Ont.			Afr.			Total Franc Min		
	L	N	%	L	N	%	L	N	%
<b>Francophones minoritaires</b>									
Enchaînements	30	52	<b>57.7</b>	27	52	<b>51.9</b>	57	104	<b>54.8</b>
Liaisons obligatoires	176	204	<b>86.3</b>	173	204	<b>84.8</b>	349	408	<b>85.5</b>
Liaisons facultatives	16	68	<b>23.5</b>	10	68	<b>14.7</b>	26	136	<b>19.1</b>
Liaisons interdites	35	56	<b>62.5</b>	40	56	<b>71.4</b>	75	112	<b>67.0</b>
<b>Total</b>	257	380	<b>67.6</b>	250	380	<b>65.7</b>	507	760	<b>66.7</b>

**Tableau 4.5. Production de liaisons et enchaînements francophones minoritaires**

Les productions des francophones majoritaires se rapprochent davantage des normes orthoépiques du français standard que celles des francophones minoritaires dans le sens où leurs productions de liaisons obligatoires sont réalisées en plus grand nombre, les liaisons interdites sont moins souvent réalisées et sont ainsi plus conformes à la règle qui demande de ne pas les réaliser (et ont en cela un plus grand pourcentage de séquences « justes »).

### **1.2.3. Les productions de liaisons obligatoires – tous francophones**

En ce qui concerne les productions des francophones (tableau 4.6.) majoritaires et minoritaires, nous remarquons que pour la grande majorité des cas, soit les liaisons obligatoires sont réalisées et enchaînées avec la bonne consonne (934/1020), soit elles ne sont pas réalisées du tout (59/1020). Il y a quelques cas d'hésitations, d'erreurs et de pataquès et quelques rares cas (2) où la consonne de liaison n'est pas enchaînée.

Le total des liaisons « correctement » réalisées et enchaînées (a) est de 934 + 10 reprises (e et i) menant à une liaison bien formée, soit un total de 944 /1020. Ainsi, si l'on considère seulement les liaisons bien formées dès la première tentative, cela représente 91.6% de liaisons réalisées et enchaînées (a), et si l'on considère toutes les liaisons réalisées et enchaînées dès la première tentative et après des reprises (a + e + i), on obtient alors 92.5%.

Les liaisons non-réalisées sont au nombre de 59/1020 et tous les autres cas représentent un total de 17 sur 1220 occurrences (1.67 %). Plus précisément, les liaisons sont réalisées dans les conditions suivantes :

- 91.6 % : correctement liée et enchaînée dès la première production ; (a)
- 5.8 % : non-réalisée ; (b)
- 0.2 % : consonne de liaison produite, mais non-enchaînée ; (c)
- 0.9 % : mal prononcée ; (d)
- 0.9 % : reprise avec bonne liaison ; (e)
- 0.1 % : reprise sans liaison ; (f)
- 0 % : reprise avec mauvaise liaison ; (g)
- 0.1 % : fausse liaison (pataqués), consonne rajoutée pour faire une liaison ; (h)
- 0.1 % : fausse liaison (pataqués) + reprise + correction (et bonne liaison) ; (i)
- 0.3 % : phonème ou syllabe du mot d'avant mal prononcé (comme dans *vieille*); (j)
- 0 % : phonème ou syllabe du mot d'après mal prononcée ; (k)
- 0.1 % : fausse liaison (pataqués) + reprise, mais PAS DE liaison ; (l)

a	%	b	%	d	%	e	%	j	%	c	%	f	%	h	%	i	%	L	%	g	%	k	%	
934	91,6	59	5,8	9	0,9	9	0,9	3	0,3	2	0,2	1	0,1	1	0,1	1	0,1	1	0,1	0	0	0	0	1020

**Tableau 4.6. Production de liaisons obligatoires tous francophones**

Notons que dans notre étude seules **deux liaisons ne sont pas enchaînées** sur 1020 occurrences (soit 0.2 %). Les résultats en ligne du projet PFC (retrouvés le 24 octobre 2010) sur les productions de francophones (tableau 4.7.) indiquent une tendance similaire,; une très faible proportion de liaisons non-enchaînées. Les Français en ont fait 51 sur 7018 sites codés (0.73%), les francophones L1 en ont fait 8 sur 2352 (0.34%) et les francophones L2 en ont fait 29 sur 1921 (1.51%).

De plus, cela prend en compte toutes les liaisons (obligatoires, facultatives et interdites), ce qui explique un plus grand nombre d'hésitations et d'épenthèses que dans notre étude, car nous analysons les catégories séparément. Or, il paraît logique qu'il y ait moins d'hésitations avec les liaisons obligatoires, dans la mesure où les mots dans les liaisons obligatoires de notre étude sont pour la plupart des mots fréquents et les structures sont également plus fréquentes.

Enfin, précisons que les séquences du projet PFC peuvent appartenir à plusieurs catégories à la fois, puisque, si on ajoute le nombre d'occurrences qui apparaissent sur le site dans les tableaux, on en obtient 2522, alors que le site indique « 2352 » sites codés, soit une

différence de 170. Cela signifie que les 170 occurrences ayant créé des hésitations appartiennent aussi aux autres catégories (non réalisées, incertaines etc.). Notre type de codage est différent de celui du projet PFC.

L. réal	%	NR	%	NE	%	inc.	%	ép.	%	hésit.	%	sites codés
1278	54,34	1058	44,98	8	0,3	5	0,21	3	0,13	170	7,23	2352

**Tableau 4.7. Résultats projet PFC, en ligne (notre tableau à partir des résultats en ligne)**

L. réal : liaisons réalisées (et enchaînées)      NR : liaisons non réalisées  
 NE : liaisons réalisées non-enchaînées      inc. : liaisons incertaines  
 ép. : liaisons épenthétiques      hésit. : hésitations

Mallet (2008) précise que ces liaisons non-enchaînées, contrairement aux politiciens d'Encrevé (1988), ne sont pas le résultat d'un effet de style, mais de reprises ou hésitations.

#### 1.2.4. Les productions de francophones et les catégories grammaticales

Bien que toutes les liaisons appartiennent à la même catégorie de « liaisons obligatoires », certaines liaisons sont réalisées dans une plus grande proportion que d'autres et semblent être plus « obligatoires » que d'autres.

Le tableau 4.8. regroupe les pourcentages obtenus pour tous les groupes francophones (majoritaires et minoritaires) par ordre décroissant de production pour chaque catégorie grammaticale. Dans notre étude, parmi les douze sous-catégories composant les liaisons obligatoires dans le groupe nominal seules trois ne sont pas réalisées à 100% : [adj. qualificatif + nom], [adj. ordinal + nom], et [nom + adj.] (dans expressions figées).

La catégorie [nom + adj.] n'est utilisée qu'avec les expressions figées. En dehors des expressions figées, c'est une liaison sinon interdite, tout au moins extrêmement rare au singulier (garçon intelligent) et une liaison facultative au pluriel (des hommes intelligents).

Parmi les contextes réalisés à moins de 97% se trouvent les trois contextes suivants :

**- Adjectif qualificatif + nom (86.6%) :** *bon ami, moyen-âge, plein air, beaux*

*Allemands, dernier étage, léger ennui, ancien édifice, grand immeuble, nouveaux*

*étudiants, vieux écrits, nouvelles idées, belles acclamations, folles aventures, faux-airs, vieux habits, vieilles éponges.*

- **Nom + adjectif (68.3%)** : *accent aigu, États-Unis, fait accompli*

- **Adjectif ordinal + nom (52.5%)** : *premier amour, second étage*

Rappelons que dans l'étude de Boula de Mareüil, Adda-Decker et Gendner (2003) la séquence [dét. + ...] est réalisée à 95.3% et la séquence [adj. + nom] est réalisée à 72.5%. Mais ces résultats prennent en compte tous les contextes possibles, y compris les cas de liaisons facultatives et interdites, tandis que nous ne considérons que les liaisons obligatoires. Dans notre étude, si l'on prend en compte tous les contextes d'adjectifs + noms, nous obtenons 675 liaisons réalisées sur 740, soit 91.22%.

Nous voyons ici une nette différence entre des situations où la liaison est réalisée par tous les francophones, et les situations où la liaison n'est pas systématique: dans les contextes [adj. qual. + nom], [nom + adj.] et [adj. ord. + nom]. Nous considérons qu'une seule occurrence non-réalisée pour un groupe ou une personne, dans une catégorie, ne représente pas une tendance. Par exemple, 7 liaisons réalisées sur 8 (bien que 87.5%) n'est pas représentatif en soi (une seule occurrence n'a pas été réalisée), comme dans les pourcentages du groupe majoritaire, ici 23/24 soit 95.8% pour le contexte [adj. num. + chiffre].

	nb	T	%
art def + nom	40	40	<b>100,0</b>
préposition + nom	20	20	<b>100,0</b>
adj dém + nom	20	20	<b>100,0</b>
adj num + nom	140	140	<b>100,0</b>
adj pos + nom	139	140	<b>99,3</b>
art indef + nom	99	100	<b>99,0</b>
prép + article	59	60	<b>98,3</b>
adj ind + nom	39	40	<b>97,5</b>
adj num + chiffre	39	40	<b>97,5</b>
adj qual + nom	277	320	<b>86,6</b>
nom + adj	41	60	<b>68,3</b>
adj ord + nom	21	40	<b>52,5</b>
<b>total et %</b>	<b>934</b>	<b>1020</b>	<b>91,6</b>

**Tableau 4.8. Liaisons obligatoires et catégories grammaticales tous francophones**

Étudions si les catégories grammaticales sont traitées de la même façon pour tous les groupes (tableau 4.9.). Nous avons séparé les productions supérieures à 95% pour lesquelles seule une occurrence n'a pas été réalisée par un des groupes, et les productions inférieures à 95% pour lesquelles plusieurs occurrences n'ont pas été réalisées dans tous les groupes.

Pour le groupe majoritaire, 7 contextes sur 12 ont été réalisés à 100%, 1 contexte à 98.8%, 1 contexte à 94.8% et 2 contextes à 77.8% et 70.8%. Toutefois si l'on considère le contexte adj [adj. qual. + nom], celui-ci est réalisé à 95% par les Européens et les Québécois, et à 93.8% par le groupe bilingue, soit 94.8% pour le groupe de francophones majoritaires. D'un côté, on peut considérer que 95% est représentatif d'une liaison obligatoire. D'un autre côté, aucun groupe ni sous-groupe n'a réalisé les liaisons de façon systématique, il y a 10 liaisons non-réalisées sur 192.

On peut donc dire que pour le groupe majoritaire, les liaisons sont « obligatoires » pour tous les contextes sauf le contexte [adj. qual. + nom], contexte quasi-obligatoire et les catégories [nom + adj.] (expressions figées) et [adj. ord. + nom].

Majoritaires	Fr + Belg.			Québec			Bil.			Tot. Fr. Maj		
	nb	T	%	nb	T	%	nb	T	%	nb	T	%
art def + nom	10	10	<b>100</b>	8	8	<b>100</b>	6	6	<b>100</b>	24	24	<b>100</b>
préposition + nom	5	5	<b>100</b>	4	4	<b>100</b>	3	3	<b>100</b>	12	12	<b>100</b>
adj dém + nom	5	5	<b>100</b>	4	4	<b>100</b>	3	3	<b>100</b>	12	12	<b>100</b>
adj num + nom	35	35	<b>100</b>	28	28	<b>100</b>	21	21	<b>100</b>	84	84	<b>100</b>
art indef + nom	25	25	<b>100</b>	20	20	<b>100</b>	15	15	<b>100</b>	60	60	<b>100</b>
prép + article	15	15	<b>100</b>	12	12	<b>100</b>	9	9	<b>100</b>	36	36	<b>100</b>
adj ind + nom	10	10	<b>100</b>	8	8	<b>100</b>	6	6	<b>100</b>	24	24	<b>100</b>
adj pos + nom	35	35	<b>100</b>	28	28	<b>100</b>	20	21	<b>95.2</b>	83	84	<b>98.8</b>
adj num + chiffre	10	10	<b>100</b>	7	8	<b>87.5</b>	6	6	<b>100</b>	23	24	<b>95.8</b>
adj qual + nom	76	80	<b>95</b>	61	64	<b>95.3</b>	45	48	<b>93.8</b>	182	192	<b>94.8</b>
nom + adj	13	15	<b>86.7</b>	8	12	<b>66.7</b>	7	9	<b>77.8</b>	28	36	<b>77.8</b>
adj ord + nom	9	10	<b>90</b>	5	8	<b>62.5</b>	3	6	<b>50</b>	17	24	<b>70.8</b>
total et %	150	151	<b>97.3</b>	119	122	<b>94.6</b>	89	93	<b>94.1</b>	358	366	<b>95.6</b>

**Tableau 4.9. Liaisons obligatoires et catégories grammaticales – francophones majoritaires**

En ce qui concerne le groupe minoritaire (tableau 4.10.), dans 6 catégories sur 12 les liaisons ont été réalisées à 100%. Trois catégories sont pratiquement réalisées à 100% (seule



une occurrence n'a pas été réalisée) : les contextes [art. ind. + nom] : 39/40 ; [prép. + art.] : 23/24 ; [adj. ind. + nom] : 15/16.

Par contre, les contextes dans lesquels les liaisons sont réalisées en plus faibles proportions sont les contextes [adj. qual. + nom] (74.2%), [nom + adj.] (54.2%) et [adj. ord. + nom] (25%). Tous francophones confondus, les liaisons dans le contexte [adj. qual. + nom] sont réalisées à 86.6%.

Les tendances des francophones minoritaires sont similaires à celles des francophones majoritaires puisque ceux-ci ont réalisé nettement moins de liaisons dans les mêmes catégories. Le pourcentage est toutefois un plus faible pour les francophones minoritaires : 74.2% vs 86.6%, 54.2% vs 68.3%, 25% vs 52.5%. Nous voyons ici que lorsque les liaisons sont réalisées en plus faibles proportions par les francophones majoritaires, elles sont réalisées en encore plus faibles proportions par les francophones minoritaires.

	Ontario			Pays africains			Total min.		
	nb	T	%	nb	T	%	nb	T	%
art def + nom	8	8	<b>100</b>	8	8	<b>100</b>	16	16	<b>100</b>
préposition + nom	4	4	<b>100</b>	4	4	<b>100</b>	8	8	<b>100</b>
adj dém + nom	4	4	<b>100</b>	4	4	<b>100</b>	8	8	<b>100</b>
adj num + nom	28	28	<b>100</b>	28	28	<b>100</b>	56	56	<b>100</b>
adj pos + nom	28	28	<b>100</b>	28	28	<b>100</b>	56	56	<b>100</b>
adj num + chiffre	8	8	<b>100</b>	8	8	<b>100</b>	16	16	<b>100</b>
art indef + nom	19	20	<b>95</b>	20	20	<b>100</b>	39	40	<b>97.5</b>
prép + article	12	12	<b>100</b>	11	12	<b>91.7</b>	23	24	<b>95.8</b>
adj ind + nom	8	8	<b>100</b>	7	8	<b>87.5</b>	15	16	<b>93.8</b>
adj qual + nom	50	64	<b>78.1</b>	45	64	<b>70.3</b>	95	128	<b>74.2</b>
nom + adj	6	12	<b>50</b>	7	12	<b>50</b>	13	24	<b>54.2</b>
adj ord + nom	1	8	<b>12.5</b>	3	8	<b>37.5</b>	4	16	<b>25</b>
total et %	119	120	<b>86.3</b>	118	204	<b>84.8</b>	237	240	<b>85.5</b>

**Tableau 4.10. Liaisons obligatoires et catégories grammaticales – francophones minoritaires**

Nous nous trouvons donc face à une question cruciale : pourquoi la liaison est-elle réalisée avec les adjectifs possessifs, numéraux et indéfinis à plus de 97%, alors qu'elle est réalisée en beaucoup plus faible proportion avec les expressions figées, les adjectifs ordinaux et qualificatifs ? Pour mieux comprendre le phénomène, penchons-nous sur ces trois cas.

Tout d’abord, les expressions figées représentent une catégorie à part qui est réalisée en plus faible proportion par tous les groupes (tableaux 4.11. et 4.12.). Nos résultats pour les francophones majoritaires (87.5%) sont proches de ceux de l’étude de Gadet (1989 : 54) qui a obtenu 85% pour l’expression figée de *haut en bas*. Dans notre étude, les francophones minoritaires réalisent les liaisons dans ce contexte en plus faible proportion (72.91%) que les francophones majoritaires (87.5%).

expres. figées	Fr + B			Québec			Bil			Ontario			Afrique			total		
	nb	T	%	nb	T	%	nb	T	%	nb	T	%	nb	T	%	nb	T	%
	27	30	<b>90</b>	20	24	<b>83</b>	16	18	<b>89</b>	17	24	<b>71</b>	18	24	75	98	120	<b>81.66</b>

**Tableau 4.11. Expressions figées - Tous francophones**

expres. figées	Fr majoritaires			Fr. minoritaires			Total		
	nb	T	%	nb	T	%	nb	T	%
	63	72	<b>87.5</b>	35	48	<b>72.91</b>	98	120	<b>81.66</b>

**Tableau 4.12. Expressions figées - Francophones majoritaires et minoritaires**

Ensuite, si l’on considère les adjectifs qualificatifs et ordinaux ainsi que les expressions figées (tableau 4.13.), on voit que, bien que *États-Unis* soit une expression figée, la séquence est correctement réalisée par tous les locuteurs (100%), alors que *accent aigu* n’est réalisé qu’à 60% et *fait accompli* à 55%. Comme nous l’avons vu au chapitre 3, le taux de cooccurrence de *fait accompli* est de 12, d’*accent aigu* est de 1, (c’est-à-dire que le taux de cooccurrence est très bas), alors que le taux de cooccurrence de *États-Unis* est + de 1000. Ceci suggère que pour les expressions figées, la fréquence pourrait jouer un rôle important.

Booij et de Jong (1987) ont remarqué que les liaisons étaient réalisées à 100% dans certaines expressions figées. Nous pensons et allons tester plus loin, que plus les expressions sont fréquentes, plus les liaisons sont réalisées dans une même catégorie grammaticale, ici les expressions figées.

On remarque le même phénomène de fréquence avec les adjectifs qualificatifs. Huit séquences contenant un adjectif qualificatif sont réalisées à 100% (par tous les

francophones), alors que d'autres sont réalisées à 95%, 90%, 80%, 75%, 70%, 60%, 55% et 30%. Ce n'est donc pas seulement la catégorie grammaticale qui joue un rôle, mais ce pourrait aussi être la fréquence d'une séquence à l'intérieur d'une catégorie ou sous-catégorie.

	Maj		Eur		Québ		min		Ont		Afr.		Bil		Tous	
	%F	%NF	%F	%NF	%F	%NF	%F	%NF	%F	%NF	%F	%NF	%F	%NF	%F	%NF
bon ami	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
moyen-âge	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
plein air	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
nouv. étud.	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
nouv.idées	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
beaux All.	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
faux-airs	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
États-Unis	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
vieux habits	100	0	100	0	100	0	100	0	100	0	100	0	100	0	100	0
vieux écrits	100	0	100	0	100	0	90	10	100	0	75	25	100	0	95	5
folles av.	100	0	100	0	100	0	90	10	100	0	75	25	100	0	95	5
belles accl.	100	0	100	0	100	0	80	20	100	0	50	50	100	0	90	10
vieilles ép.	100	0	100	0	100	0	80	20	100	0	50	50	100	0	90	10
dernier étage	100	0	100	0	100	0	60	40	50	50	50	50	100	0	80	20
ancien édifice	100	0	100	0	100	0	60	40	75	25	25	75	100	0	80	20
second étage	90	10	100	0	80	20	70	30	75	25	50	50	100	0	80	20
grand im.	100	0	100	0	100	0	50	50	25	75	50	50	100	0	75	25
prem.amour	90	10	100	0	80	20	50	50	25	75	50	50	100	0	70	30
accent aigu	70	30	60	40	80	20	50	50	50	50	50	50	50	50	60	40
fait accompli	80	20	100	0	60	40	30	70	0	100	25	75	100	0	55	45
léger ennui	40	60	40	60	40	60	20	80	0	100	25	75	50	50	30	70
<b>total</b>	<b>97,5</b>	<b>2,5</b>	<b>98,1</b>	<b>1,96</b>	<b>96,9</b>	<b>3,1</b>	<b>90,2</b>	<b>9,8</b>	<b>89,7</b>	<b>10,3</b>	<b>86,8</b>	<b>13,2</b>	<b>98,1</b>	<b>1,9</b>	<b>93,8</b>	<b>6,2</b>

**Tableau 4.13. Expressions figées et adjectifs, liaisons obligatoires - tous francophones**

Notre hypothèse est que les séquences pour lesquelles les liaisons sont le moins réalisées à l'intérieur d'une même catégorie grammaticale sont les séquences dans lesquelles les mots sont moins fréquents et pour lesquelles le taux de cooccurrence est plus faible. Pour vérifier notre hypothèse du rôle de la fréquence, nous devons déterminer les différents types de fréquence, la fréquence des items, et ainsi étudier le lien entre la fréquence et la production de liaison.

### **1.2.5. Les productions de francophones et la fréquence lexicale**

Pour déterminer le rôle de la fréquence, nous avons décidé d'utiliser plusieurs banques de données, afin de nous assurer de la validité de nos hypothèses quant au rôle de la fréquence.

#### **1.2.5.1. Analyse du mot1 avec *Vocabprofil* et *Lexique***

Tout d'abord, si l'on considère la banque de données de *Vocabprofil*, le lien entre la fréquence des mots et la réalisation des liaisons est particulièrement frappant. Le Mot1 représente la fréquence de la catégorie syntaxique (si le mot1 est un article, la catégorie syntaxique est probablement article + nom, bien qu'elle puisse aussi être article + adjectif), mais aussi dans une moindre mesure la fréquence du mot1 tandis que le mot2 représente principalement la fréquence lexicale du mot2.

La raison pour laquelle il n'y a pas de mots Hors-Liste (dorénavant HL) pour les Mots1 dans le texte est parce que les mots1 sont principalement des articles, prépositions, et autres mots très fréquents. Par contre, puisque le Mot2 représente davantage la fréquence lexicale, on trouve dans les mots2 des mots plus rares appartenant à la catégorie HL.

Comme on le voit dans les tableaux 4.14. et 4.15., plus les mots sont fréquents, plus le taux de réalisation est élevé pour tous les locuteurs francophones. Les liaisons avec les mots1 appartenant à la catégorie K1 (les 1000 mots les plus fréquents) sont réalisées à 97.4% par les francophones majoritaires et à 87.8% par les francophones minoritaires. Les mots appartenant à la catégorie K2 sont réalisés à 70.8% par les francophones majoritaires et à 50% par les francophones minoritaires. Les mots appartenant à la catégorie K3 sont réalisés entre à 58.3% par les francophones majoritaires et à 50% par les francophones minoritaires. Il n'y a pas de mots1 appartenant à la catégorie H.L.

Il apparaît donc que la fréquence du mot1 joue un rôle important dans la production de liaisons. Étant donné que le mot1 indique en partie la fréquence d'une structure grammaticale, ce qu'Ellis (2002) et Bybee (2001) appellent *type frequency*, la fréquence de la structure grammaticale, où le type de structure est un indicateur fiable de production de liaison.

mot1	Fr.			B.			QB			Bil			Fr. maj		
	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
3	190	192	99	48	48	100	185	192	96.4	138	144	95.8	561	576	97.4
2	6	8	75	1	2	50	6	8	75	4	6	66.7	17	24	70.8
1	3	4	75	0	1	0	2	4	50	2	3	66.7	7	12	58.3
total	199	204	97.5	49	51	96.1	193	11	94.60	144	153	94.10	585	612	95.6

Tableau 4.14. Fréquence, production liaisons obligatoires, mot1 *Vocabprofil*, francophones majoritaires

mot1	On			Af.			Fr. min			ts		
	L	N	%	L	N	%	L	N	%	L	N	%
3	170	192	88.5	167	192	87	337	384	87.8	898	960	93.5
2	4	8	50	4	8	50	8	16	50	25	40	62.5
1	2	4	50	2	4	50	4	8	50	11	20	55
total	176	204	86.30	173	204	84.80	349	408	85.5	934	1020	91.60

Tableau 4.15. Fréquence, production liaisons obligatoires, mot1 *Vocabprofil*, francophones minoritaires

Si l'on examine la fréquence du mot1 avec la banque de données de *Lexique* (tableaux 4.16., 4.17. et 4.18.), on obtient des tendances similaires : plus les mots sont fréquents, plus les liaisons sont réalisées. Toutefois les chiffres sont encore plus représentatifs du rôle de la fréquence avec cette banque de données. Les liaisons sont réalisées à 99.17% par tous les francophones avec les mots1 les plus fréquents (93.5% avec *Vocabprofil*), à 91.19% avec les mots moyennement fréquents (62.5% avec *Vocabprofil*), et à 80.83% avec les mots les moins fréquents (à 55% avec *Vocabprofil*).

Les francophones majoritaires réalisent les liaisons avec les mots les plus fréquents à 99.54% (97.4% avec *Vocabprofil*), à 98.02% avec les mots moyennement fréquents (70.8% avec *Vocabprofil*), et à 85.42% avec les mots les moins fréquents (58.3%).

Les francophones minoritaires réalisent les liaisons avec les mots les plus fréquents à 98.61% (87.8% avec *Vocabprofil*), à 80.95% avec les mots moyennement fréquents (50% avec *Vocabprofil*), et à 70.83% avec les mots les moins fréquents (50% avec *Vocabprofil*).

L'analyse des résultats avec *Lexique* suggère les tendances suivantes : avec les mots les plus fréquents (501-8000), il y a peu de différences entre les groupes majoritaires : 99.54% (tableau 4.17) et minoritaires : 98.61% (tableau 4.18).

Les différences sont les plus visibles avec les mots moyennement fréquents et rares.

Les francophones majoritaires produisent 98.02% de liaisons avec les mots moyennement fréquents (mots à fréquence 41-500 dans tableau 4.17.), alors que les francophones minoritaires en produisent 80.95% (tableau 4.18). Enfin, avec les mots les plus rares (0 à 40), les francophones majoritaires produisent 85.42% de liaisons, alors que les francophones minoritaires en produisent 73.96%.

Tous les groupes font plus de liaisons avec les mots très fréquents qu'avec les mots moyennement fréquents, et ils en font plus avec les mots moyennement fréquents qu'avec les mots plus rares. C'est aussi dans cette catégorie que des différences se remarquent à l'intérieur des groupes. Les Français produisent 91.67% de liaisons avec les mots plus rares, les Québécois en produisent 83.33%, et les personnes ayant un parent francophone en produisent 80.56% alors que pour les mots les plus fréquents il n'y avait pas de différences (seule une occurrence différenciait un groupe de tous les autres groupes majoritaires).

Fréq Mot1	c	Ts		
		L	N	%
0 à 40	abc	194	240	<b>80.83</b>
41 à 500	defgh	383	420	<b>91.19</b>
501-8000	ijklm	357	360	<b>99.17</b>
Total		934	1020	<b>91.6</b>

Tableau 4.16. Fréquence Mot1 Lexique liaisons obligatoires – Tous francophones

Fréquence	Mot1	francophones majoritaires														
		Fr.			B.			QB			1PF			Maj		
Lexique	c	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
0 à 40	abc	44	48	<b>91.67</b>	10	12	<b>83.33</b>	40	48	<b>83.33</b>	29	36	<b>80.56</b>	123	144	<b>85.42</b>
41 à 500	defgh	83	84	<b>98.81</b>	21	21	<b>100.00</b>	81	84	<b>96.43</b>	62	63	<b>98.41</b>	247	252	<b>98.02</b>
501-8000	ijklm	72	72	<b>100.00</b>	18	18	<b>100.00</b>	72	72	<b>100.00</b>	53	54	<b>98.15</b>	215	216	<b>99.54</b>
Total		199	204	<b>97.55</b>	49	51	<b>96.08</b>	193	204	<b>94.61</b>	144	153	<b>94.12</b>	585	612	<b>95.59</b>

Tableau 4.17. Fréquence Mot1 Lexique liaisons obligatoires – Francophones majoritaires

Fréquence		francophones minoritaires								
Mot1		On			Af.			Ts		
Lexique	c	L	N	%	L	N	%	L	N	%
0 à 40	abc	37	48	<b>77.08</b>	34	48	<b>70.83</b>	71	96	<b>73.96</b>
41 à 500	defgh	68	84	<b>80.95</b>	68	84	<b>80.95</b>	136	168	<b>80.95</b>
501-8000	ijklm	71	72	<b>98.61</b>	71	72	<b>98.61</b>	142	144	<b>98.61</b>
Total		176	204	<b>86.27</b>	173	204	<b>84.80</b>	349	408	<b>85.54</b>

**Tableau 4.18. Fréquence Mot1 Lexique liaisons obligatoires – Francophones minoritaires**

Il apparaît donc que, quelle que soit la banque de données utilisée, la fréquence des mots1 joue un rôle indéniable sur la production de liaisons des francophones majoritaires et minoritaires. Lorsque les mots1 sont très fréquents, tous les francophones produisent la liaison (142/144). Avec les mots moyennement fréquents, les francophones majoritaires réalisent encore presque toutes les liaisons (247/252), mais lorsque les mots sont moins fréquents, ils les produisent en plus faible quantité, bien que 85.42% soit tout de même une grande proportion.

Les francophones majoritaires produisent plus de liaisons (85.42%) avec les mots les plus rares que les francophones minoritaires n'en produisent avec les mots moyennement fréquents (80.95%). Ceci s'explique de la façon suivante. Les francophones majoritaires ont sans doute entendu les mots les plus rares plus souvent que les francophones minoritaires n'ont entendu les mots moyennement fréquents puisque les francophones minoritaires de notre étude étaient en contact restreint avec le français.

On peut présumer que les francophones minoritaires ont entendu les mots les plus fréquents suffisamment de fois (même si le nombre est certainement plus bas que pour les francophones majoritaires) pour que les séquences soient ancrées dans leur lexique. Cela nous amène à conclure que le rôle de la fréquence du mot1 pourrait être un indice fiable de productions de liaisons, ce que nous allons vérifier au moyen de l'analyse à régression avec *GoldVarb*.

### 1.2.5.2. Analyse du mot2 avec *Vocabprofil* et *Lexique*

En ce qui concerne les liaisons avec les mots2 (tableaux 4.19. et 4.20.) avec *Vocabprofil*, nous obtenons des tendances similaires à celles obtenues avec les mots1. Globalement, plus les mots2 sont fréquents, plus les liaisons sont réalisées par tous les groupes.

Les liaisons avec les mots appartenant à la catégorie K1 sont réalisées à 97.75% par les francophones majoritaires et à 91.6% par les francophones minoritaires. Les liaisons avec les mots appartenant à la catégorie K2 sont réalisées à 97.22% par les francophones majoritaires et à 83.3% par les francophones minoritaires.

Les francophones majoritaires suivent l'ordre de fréquence : plus les mots sont fréquents, plus ils réalisent de liaisons (97.75% pour les mots les plus fréquents, 97.22% pour les mots moyennement fréquents, 91.67% pour les mots plus rares, et 81.94% pour les mots rares). Par contre, les francophones minoritaires ne suivent pas exactement cette tendance : ils réalisent une plus grande proportion de liaisons avec les mots HL (68.8%), (contre 81.94% pour les francophones majoritaires) qu'avec les mots K3 (31.3%) (contre 91.67% pour les francophones majoritaires).

Il existe aussi des différences entre les groupes majoritaires et minoritaires pour les mots les plus fréquent (91.6% vs 97.75%), mais ce sont surtout avec les mots moyennement fréquents (K2) peu fréquents (K3) et rares (HL) que les différences sont les plus visibles.

On peut aussi penser que pour les francophones majoritaires, la fréquence lexicale (mot2) est (presque) aussi importante que la fréquence syntaxique (mot1), alors que pour les francophones minoritaires, la fréquence syntaxique semble plus importante que la fréquence lexicale, ce qui pourrait s'expliquer par le fait qu'ils sont moins familiers avec les mots moins fréquents et devraient se fier à des indices généraux (structures) plutôt qu'à des indices particuliers (les items).



La fréquence du mot2 semble alors jouer un rôle important, mais un rôle moins important que le mot1 avec la banque de données de *Vocabprofil*. Cela suggère que la fréquence de la structure syntaxique pourrait être un élément plus important que la fréquence lexicale (du mot2).

mot2	Fr.			B.			QB			Bil			Fr maj		
	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
3	146	148	98.6	37	37	100	143	148	96.6	108	111	97.3	434	444	97.75
2	24	24	100	6	6	100	24	24	100	16	18	88.9	70	72	97.22
1	8	8	100	2	2	100	6	8	75	6	6	100	22	24	91.67
0	21	24	87.5	4	6	66.7	20	24	83.3	14	18	77.8	59	72	81.94
total	199	204	97.5	49	51	96.1	193	204	94.6	144	153	94.1	585	612	95.59

Tableau 4.19. Fréquence, production liaisons obligatoires, mot2 *Vocabprofil* - francophones majoritaires

mot2	On			Af.			Fr min			ts		
	L	N	%	L	N	%	L	N	%	L	N	%
3	135	148	91.2	136	148	91.9	271	296	91.6	705	740	95.3
2	20	24	83.3	20	24	83.3	40	48	83.3	110	120	91.7
1	3	8	37.1	2	8	25	5	16	31.3	27	40	67.5
0	18	24	75	15	24	62.5	33	48	68.8	92	120	76.7
total	176	204	86.30	173	204	84.80	349	408	85.5	934	1020	91.6

Tableau 4.20. Fréquence, production liaisons obligatoires, mot2 *Vocabprofil* - francophones minoritaires

Après avoir calculé la fréquence des mots2 et la production de liaisons selon les catégories de fréquence avec *Vocabprofil*, calculons-les à présent avec *Lexique* pour déterminer si on observe les mêmes tendances.

Les séparations en catégories de fréquence a, b, c, d, e, f, etc. ont au départ été faites en les divisant en 13 parties qui semblaient représenter les différents niveaux de fréquence. Toutefois, pour l'analyse, il paraissait plus pertinent de les séparer en trois catégories : les mots les plus fréquents, moyennement fréquents, mots rares (catégorisation plus similaire à *Vocabprofil*).

Si l'on ajoute les chiffres obtenus avec *Lexique* des groupes très rares (a + b), (tableaux 4.21., 4.22. et 4.23.), ceux des groupes moyennement fréquents (c + d + e) et ceux des groupes fréquents (f + g + h + i), on observe les mêmes tendances qu'avec *Vocabprofil* et

de façon plus flagrante. Comme il n'y a eu aucun résultat dans la fréquence f entre 80 et 101, nous n'avons pas inclus f dans les tableaux 4.21-4.23.

Les tendances sont les mêmes en ce qui concerne la fréquence du mot2 pour les liaisons obligatoires. Les liaisons sont pratiquement toutes réalisées avec les mots les plus fréquents par tous les francophones (97.6%), sont faites dans la majorité des cas avec les mots moyennement fréquents (88.8%) et en plus faible proportion avec les mots plus rares (82.9%). Les tendances sont les mêmes avec le mot2 qu'avec le mot1.

Les francophones majoritaires réalisent les liaisons avec les mots2 les plus fréquents à 99.1%, les liaisons avec les mots2 moyennement fréquents à 95.1% et les liaisons avec les mots les plus rares à 89.9%. Les francophones minoritaires réalisent les liaisons avec les mots2 les plus fréquents à 95.4%, les liaisons dans les mots2 moyennement fréquents à 79.2% et les liaisons avec les mots les plus rares à 72.3%.

Ceci indique que la fréquence du mot2 semble jouer un rôle important dans la production de liaison pour tous les francophones. Tous les groupes font plus de liaisons avec les mots les plus fréquents qu'avec les mots moyennement fréquents, et plus de liaisons avec les mots moyennement fréquents qu'avec les mots plus rares.

C'est également dans la catégorie des mots rares que des différences se remarquent même à l'intérieur des groupes. Les Français produisent 92.9% de liaisons avec les mots plus rares, alors que les Québécois en produisent 89.3% et les personnes bilingues 88.1%, alors que pour les mots les plus fréquents, seules deux occurrences différenciaient un groupe de tous les autres groupes majoritaires.

C'est avec les mots moyennement fréquents et les mots les plus rares que les différences entre les francophones majoritaires et minoritaires sont les plus flagrantes : 89.9% pour les francophones majoritaires, et 72.3% pour les francophones minoritaires avec les mots les plus rares.

Fréq Mot2		tous franc.		
Lexique	c	L	N	%
0 - 20	ab	232	280	82,9
21 – 80*	cde	213	240	88,8
101 et plus	ghi	371	380	97,6
Total		816	900	90,7

**Tableau 4.21. Fréquence Mot2 Lexique liaisons obligatoires – Tous francophones**

\* il n'y a eu aucun résultat avec f (fréquence entre 80 et 101)

Fréquence		francophones majoritaires														
Mot2		Fr			B.			QB			1PF			maj		
Lexique	c	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
0 - 20	ab	52	56	92,9	12	14	85,7	50	56	89,3	37	42	88,1	151	168	89,9
21 – 80*	cde	47	48	97,9	12	12	100	45	48	93,8	33	36	91,7	137	144	95,1
101 et plus	ghi	76	76	100	19	19	100	74	76	97,4	57	57	100	226	228	99,1
Total		175	180	97,2	43	45	96,1	169	180	93,9	127	135	94,1	514	540	95,2

**Tableau 4.22. Fréquence Mot2 Lexique liaisons obligatoires – Francophones majoritaires**

\* il n'y a eu aucun résultat avec f (fréquence entre 80 et 101)

Fréquence		francophones minoritaires								
Mot2		On			Af.			min		
Lexique	c	L	N	%	L	N	%	L	N	%
0 - 20	ab	43	56	76,8	38	56	67,9	81	112	72,3
21 – 80*	cde	36	48	75	40	48	83,3	76	96	79,2
101 et plus	ghi	73	76	96,1	72	76	94,7	145	152	95,4
Total		152	180	84,4	150	180	83,3	302	360	83,9

**Tableau 4.23. Fréquence Mot2 Lexique liaisons obligatoires – Francophones minoritaires**

\* il n'y a eu aucun résultat avec f (fréquence entre 80 et 101)

Lorsqu'on utilise la banque de données de *Lexique*, les productions de liaisons, selon la fréquence du mot2, suivent avec plus de précision le modèle de fréquence, que si on utilise *Vocabprofil*, et ce, pour les deux groupes de francophones. Ainsi, la fréquence du mot2 en utilisant *Lexique* semble être plus fiable pour prédire et expliquer les productions de liaisons qu'avec *Vocabprofil*, probablement parce que la banque de données de *Lexique* permet de déterminer avec plus de précision la fréquence des mots.

### 1.2.5.3. Analyse de cooccurrence mots1-mots2 avec *Lexique*

En ce qui concerne la fréquence de cooccurrence des mots1 et mots2, on observe les mêmes tendances (tableaux 4.24., 4.25. et 4.26.) qu'avec les mots1 et les mots2 analysés

séparément. Tous les francophones réalisent un plus grand pourcentage de liaisons avec les mots au taux de cooccurrence le plus élevé (98.64%) qu'avec les mots au taux de cooccurrence moyennement élevé (95.33%) et plus de liaisons avec les mots au taux de cooccurrence moyennement élevé qu'avec les mots au taux de cooccurrence plus faible (86.2%). Les différences entre les groupes majoritaires et minoritaires sont faibles avec les mots au taux de cooccurrence le plus élevé (99.24% vs 97.7%), alors qu'elles sont plus grandes avec ceux au taux moyennement fréquent (98.89% vs 90%) et encore plus visibles avec ceux au taux le plus bas (92% vs 77.5%).

La fréquence de cooccurrence des mots semble encore plus révélatrice que la fréquence des mots<sup>1</sup> et la fréquence des mots<sup>2</sup>, puisque une seule liaison sur 132 occurrences n'a pas été réalisée pour tous les groupes francophones majoritaires dans les séquences au taux de cooccurrence le plus élevé. Pour trois groupes francophones majoritaires la liaison est réalisée à 100% quand le taux de cooccurrence est supérieur à 500. Pour le groupe ayant un parent francophone, une seule liaison sur 33 n'a pas été réalisée.

Les productions des Africains et des Franco-ontariens indiquent que leurs résultats sont similaires dans toutes les catégories. Seules 3 occurrences sur 204 les différencient dans une catégorie, les 2 autres catégories ayant exactement le même nombre d'occurrences. La fréquence de cooccurrence est donc un facteur déterminant dans la production de liaisons de tous les groupes.

Fréquence				
Fréq cooc.		tous franc.		
Lexique	c	L	N	%
0 - 59	ABCD	431	500	<b>86,2</b>
60-449	EFGHI	286	300	<b>95,33</b>
500 et plus	KLM	217	220	<b>98,64</b>
Total	Total	934	1020	<b>91,57</b>

**Tableau 4.24. Fréquence de cooccurrence Lexique liaisons obligatoires – Tous francophones**

Fréquence	cooccurrence	Franc.maj														
		Fr.			B			QB			1 parent Franc.			maj		
Lexique	c	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
0 - 59	ABCD	95	100	<b>95</b>	23	25	<b>92</b>	91	100	<b>91</b>	67	75	<b>89,33</b>	276	300	<b>92</b>
60-449	EFGHI	60	60	<b>100</b>	15	15	<b>100</b>	58	60	<b>96,67</b>	45	45	<b>100</b>	178	180	<b>98,89</b>
500 et plus	KLM	44	44	<b>100</b>	11	11	<b>100</b>	44	44	<b>100</b>	32	33	<b>96,97</b>	131	132	<b>99,24</b>
Total	Total	199	204	<b>97,55</b>	49	51	<b>96,08</b>	193	204	<b>94,61</b>	144	153	<b>94,12</b>	585	612	<b>95,59</b>

Tableau 4.25. Fréquence de cooccurrence Lexique liaisons obligatoires – Francophones majoritaires

Fréquence	cooccurrence	Franc.min								
		On			Afr.			min		
Lexique	code	L	N	%	L	N	%	L	N	%
0 - 59	ABCD	79	100	<b>79</b>	76	100	<b>76</b>	155	200	<b>77,5</b>
60-449	EFGHI	54	60	<b>90</b>	54	60	<b>90</b>	108	120	<b>90</b>
500 et plus	KLM	43	44	<b>97,73</b>	43	44	<b>97,73</b>	86	88	<b>97,73</b>
Total	Total	176	204	<b>86,27</b>	173	204	<b>84,8</b>	349	408	<b>85,54</b>

Tableau 4.26. Fréquence de cooccurrence Lexique liaisons obligatoires – Francophones minoritaires

La fréquence du mot1, la fréquence du mot2 et la fréquence de cooccurrence des mots1 et mots2 semblent être des facteurs clés dans la production de liaison, facteurs qui peuvent déterminer en grande partie si une liaison sera réalisée ou non, au moins en ce qui concerne les liaisons obligatoires.

### 1.2.6. Les productions de francophones et les consonnes de liaison

Nous avons déterminé précédemment le rôle du facteur syntaxique et lexical. Voyons à présent le rôle de la consonne de liaison dans la production de liaisons obligatoires.

Les résultats en ligne du projet PFC (tableau 4.27.) indiquent que les liaisons sont réalisées par ordre décroissant avec les consonnes /z/ (43.47%), /n/ (28.93%), /t/ (26.44%), /v/ (1.4%), /R/ (0.23%). Certaines consonnes sont plus fréquemment en position de site possible. Par exemple, les cas de liaisons en /R/ sont les sites les plus rares, vu que peu de mots se trouvent dans ces sites (verbes à l'infinitif et quelques adjectifs). Les sites possibles en /v/ sont également très rares. Les sites les plus fréquents sont ceux en /z/ et en /n/.

De plus, les données ci-dessous considèrent tous les sites possibles, ce qui comprend liaisons obligatoires, facultatives et interdites. Or, si l'on considère par exemple le /n/ en liaison obligatoire, cela va restreindre les sites à la catégorie [dét. sing. + nom] (*un enfant*), et [adj. + nom] (*certain âge*), ce qui n'est pas du tout la même chose, si l'on considère les liaisons facultatives. Ceci explique pourquoi nous obtenons des résultats différents.

consonne	PFC	
	N	%
v	18	1,4
n	372	28,93
z	559	43,47
t	340	26,44
r	3	0,23
tot.	1292	

Tableau 4.27. Productions francophones L1 PFC

Dans notre étude, les consonnes de liaison jouent un rôle important dans la production des liaisons. Les liaisons obligatoires sont réalisées à 100% par les francophones majoritaires avec les consonnes /v/ et /n/, en très grande proportion avec /z/ (98.5%), en plus faible proportion avec /t/ (83.3%) et enfin, c'est avec le /R/ qu'elles sont le moins réalisées (77.8%).

Par contre, pour les francophones minoritaires, la tendance n'est pas exactement la même. Si la liaison est également réalisée à 100% avec le /v/, elle est réalisée en plus faible proportion avec /n/ (92.7%) et surtout, elle est réalisée en plus grande proportion avec /z/ (95.5%) qu'avec /n/ (92.7%). Il se peut que ce soit le cas puisque les sites de liaisons possibles sont plus fréquents avec /z/, comme indiqué dans le tableau 4.28., et que les francophones minoritaires soient davantage influencés par la fréquence de ces consonnes de liaison que les francophones majoritaires.

Enfin, c'est avec les consonnes de liaison les moins fréquentes, que l'on voit les plus grandes différences. Alors que les francophones majoritaires produisent 83.5% de liaisons avec /t/, les francophones minoritaires en produisent 57.1%. Les francophones majoritaires

produisent 77.8% de liaisons avec /R/ alors que les francophones minoritaires en produisent 25% soit le tiers.

Là aussi, on observe des tendances similaires à travers les groupes : lorsque les liaisons sont réalisées en plus faible proportion par les francophones majoritaires (avec le /t/ et le /R/), elles sont réalisées en encore plus faible proportions par les francophones minoritaires. Tous les groupes sont sensibles à la fréquence des consonnes de liaison, mais lorsque les consonnes sont rares, les francophones minoritaires produisent une très faible quantité de liaisons (25% avec /R/ pour les groupes minoritaires).

	tous francophones								
consonne	franc. Maj			franc. Min			total		
liaison LO	N		%	N		%	N		%
<b>v</b>	12	12	<b>100.0</b>	8	8	<b>100.0</b>	20	20	<b>100.0</b>
<b>n</b>	144	144	<b>100.0</b>	89	96	<b>92.7</b>	233	240	<b>97.1</b>
<b>z</b>	331	336	<b>98.5</b>	214	224	<b>95.5</b>	545	560	<b>97.3</b>
<b>t</b>	70	84	<b>83.3</b>	32	56	<b>57.1</b>	102	140	<b>72.9</b>
<b>r</b>	28	36	<b>77.8</b>	6	24	<b>25.0</b>	34	60	<b>56.7</b>
<b>tot.</b>	585	612	<b>95.6</b>	349	408	<b>85.5</b>	934	1020	<b>91.6</b>

Tableau 4.28. Consonnes de liaison et productions de liaisons obligatoires (notre étude)

## Discussion sur les résultats des francophones

D'après les résultats de cette étude, voici ce que nous pouvons déduire des hypothèses de départ. Comme on pouvait s'y attendre d'après le MBU, la fréquence lexicale influence l'acquisition et la production de liaisons. Les francophones majoritaires produisent un plus grand nombre de liaisons obligatoires et liaisons facultatives et un plus grand nombre d'enchaînements que les francophones minoritaires.

La fréquence lexicale joue un rôle déterminant à trois niveaux : la fréquence du mot1, la fréquence du mot2 et la fréquence de cooccurrence du mot1 et du mot2. La banque de

données de *Lexique* semble être plus utile et plus fiable pour ce type de recherches que la banque de données de *Vocabprofil*. Rappelons que les données utilisées dans *Lexique* ont été élaborées à partir du français oral (d'après les sous-titres de films), alors que les données de *Vocabprofil* ont été élaborées à partir de la fréquence relative des mots les uns par rapport aux autres dans des corpus écrits. Ainsi, la liaison pourrait être plus liée aux séquences entendues, qu'aux séquences lues, même pour un exercice de lecture à voix haute.

Le modèle reçu en début d'acquisition joue un rôle important, comme c'est le cas dans les études sur l'acquisition des liaisons chez les enfants francophones (Chevrot, Dugua et al., 2005). Les deux francophones bilingues qui ont une mère française ont produit légèrement plus de liaisons que les Québécois, alors qu'ils ont reçu moins d'input qu'eux, puisque le père parlait une autre langue et qu'ils n'étaient pas à 100% en contact avec le français dans les activités quotidiennes en dehors de la famille et de l'école puisque les autres activités étaient en anglais.

Nous supposons (mais ne pouvons le prouver ici vu le nombre trop restreint de locuteurs) que les mères françaises de ces deux sujets ont produit un plus grand nombre de liaisons que les mères québécoises, ce qui a eu des répercussions sur la production des sujets bilingues. C'est ce qu'on remarque dans les productions globales des Français et des Québécois. Comment expliquer autrement leur production un peu plus élevée ? Il est peu probable que ce soit à l'école qu'ils aient appris ces liaisons, sinon les locuteurs québécois les auraient également apprises à l'école. Cela signifie que la qualité de l'input dans les premières années (passées principalement avec la mère), en début d'acquisition, pourrait jouer un rôle plus important que l'input reçu plus tard, ce qu'on voit avec les bilingues de mère francophone qui ont ensuite reçu un input partagé entre deux langues, donc plus restreint que les francophones majoritaires.



Les liaisons obligatoires les plus fréquentes ont été réalisées par tous les francophones et les autres liaisons obligatoires moins fréquentes ont été réalisées en plus faibles proportions (avec les expressions figées et certains adjectifs). Pour les séquences moins fréquentes, la production de liaisons dépend des facteurs linguistiques propres à la séquence : fréquence du mot1, fréquence du mot2, fréquence de cooccurrence des mots et consonne de liaison. Et elle dépend aussi des facteurs sociaux propres aux locuteurs.

Les liaisons avec les mots les plus fréquents ont été réalisées à plus de 95% par tous les francophones, tandis que les liaisons avec les mots moins fréquents ont été réalisées à 89% par les francophones majoritaires et à 72% par les francophones minoritaires. Les tendances sont les mêmes pour tous les groupes, mais avec un taux de liaisons obligatoires et facultatives plus bas et un taux de liaisons interdites plus haut chez les francophones minoritaires.

Les francophones bilingues ont produit autant de liaisons que les francophones majoritaires, et un peu moins de liaisons facultatives et/ou enchaînements que les francophones majoritaires, mais plus de liaisons obligatoires et facultatives que les francophones minoritaires. Les liaisons interdites et les liaisons qui sont des exceptions aux règles (liaisons dans les expressions figées) ont causé plus de difficultés aux francophones minoritaires, pour les mêmes raisons que les liaisons obligatoires et facultatives ont été moins élevées dans ces groupes.

La fréquence des items joue un plus grand rôle chez les francophones minoritaires que chez les francophones majoritaires. C'est surtout pour les liaisons moins fréquentes que les différences sont les plus visibles. En effet, puisque passé un seuil de fréquence, les séquences sont acquises dans le lexique, les plus fréquentes ont été « enregistrées » dans le lexique, mais les moins fréquentes n'ont pas été entendues suffisamment de fois pour être

enregistrées. Il est donc logique que moins les séquences sont fréquentes, moins elles soient produites. La fréquence de réception est directement liée à la fréquence de production.

Les facteurs syntaxiques, phonétiques et fréquentiels nous amènent à reconsidérer les classifications des liaisons obligatoires. Contrairement à la classification de l'Académie Française, la même classification utilisée par Encrevé (1988) et Delattre (1947), on ne peut pas dire que la réalisation de la liaison soit absolue pour deux catégories de « liaisons obligatoires », les liaisons avec les expressions figées et dans le contexte [adj. + nom].

Il convient donc de préciser les catégorisations en question en rajoutant que les liaisons se font de manière systématique dans le contexte [dét. + nom], [prép. monos. (*dans*) + GN], pour quelques expressions figées les plus fréquentes (qu'il reste à déterminer) et pour la plupart des adjectifs courts et très fréquents + noms. À cela il faut préciser que le taux de fréquence des mots<sup>1</sup> et des mots<sup>2</sup> et le taux de cooccurrence des mots<sup>1</sup> et mots<sup>2</sup> doit être élevé pour que la liaison soit produite.

Notre catégorisation est alors la suivante. Les liaisons se font de façon systématique :

- Entre un déterminant et un nom ;
- Avec les adjectifs possessifs, démonstratifs et cardinaux
- Avec certaines expressions figées les plus fréquentes et mots composés,
- Avec certains adjectifs courts antéposés les plus fréquents,

Elles se feraient aussi (d'après les recherches sur les productions de liaisons de francophones précédemment citées) dans les contextes suivants de façon systématique, mais cela n'a pas été testé dans cette étude :

- Avant ou après un pronom,
- Après une préposition monosyllabique très fréquente (*dans*),
- Après un adverbe monosyllabique fréquent (*très*).

Les francophones minoritaires ont fait moins de liaisons dans toutes les catégories, sauf pour les liaisons avec les mots les plus fréquents où ils en ont fait pratiquement autant que les francophones majoritaires, et avec les liaisons interdites où ils en ont fait plus. Ceci était attendu.

D'une part, la fréquence d'écoute joue un rôle déterminant : ayant sans doute entendu moins souvent certaines liaisons ou certains contextes (qui constituent des liaisons interdites ou facultatives), les francophones minoritaires réalisent moins de liaisons que les francophones majoritaires, et le nombre de liaisons réalisées correspond au nombre de liaisons ou de contextes entendus.

D'autre part, le rôle de l'input est déterminant et se rajoute au rôle de la fréquence. Si les personnes desquelles ils ont appris le français parlent d'une certaine façon (par exemple, en faisant moins de liaisons et enchaînements), les sujets entendront moins de contextes et en produiront moins. Ceci est particulièrement visible lorsque des personnes ont un ou deux parents européens (les francophones bilingues) et ont pourtant vécu en milieu anglophone. Même s'ils sont en situation de francophone minoritaire, ils ont entendu le français de parents européens (contenant un plus grand nombre de liaisons que le français d'autres groupes francophones) et en font donc plus eux-mêmes, malgré un input par la suite restreint.

Autrement dit, cela rejoint les études portant sur les enfants francophones. L'influence des parents et de la langue parlée par les parents est un facteur important dans l'acquisition de certains traits tels que les liaisons (et enchaînements). C'est la raison pour laquelle, lorsque le français a été appris à l'école, ce qui est le cas pour les Africains, leur production est plus basse que la production des Franco-ontariens.

Par contre, si les sujets sont en contact permanent suffisamment longtemps avec une communauté linguistique qui produit un plus grand nombre de liaisons, ils ont plus de chances de produire un plus grand nombre de liaisons que s'ils étaient restés dans la même

communauté linguistique qui en fait moins. C'est ce qu'on peut remarquer avec le sujet africain FA1 qui a eu plus de contacts en Afrique avec des Français, et qui a aussi passé plusieurs années en France. Celui-ci a produit un total de 79% liaisons, ce qui est plus que la moyenne des Québécois (76.6) et très proche de la moyenne des francophones européens (79.2).

Les seules liaisons qui n'ont pas été réalisées à 100% sont pour les trois adjectifs *second étage*, *premier amour*, *léger ennui* et pour les deux expressions figées *accent aigu*, *fait accompli*. Dans un même contexte syntaxique [adj. + nom] et phonétique (/R/ + voyelle ou d qui se prononce /t/ + voyelle), les résultats diffèrent.

Nous pouvons expliquer les différences de production des francophones par l'effet de fréquence de cooccurrence de ces mots. Le contexte « second étage » n'est pas très fréquent dans la mesure où, on utilise le mot *second* quand il n'y en a pas d'autres après (il n'y a que deux items, le premier et le second). Comme il est rare qu'un immeuble n'ait que deux étages, on emploierait plutôt le mot *deuxième* et on dirait donc plus naturellement *deuxième étage* que *second étage*. Par contre, non seulement l'expression *dernier étage* est fréquente, elle est aussi très utilisée, presque lexicalisée. C'est donc la fréquence de cooccurrence qui influe le plus sur la production de liaison.

En ce qui concerne le cas de *léger ennui* qui a obtenu le pourcentage le plus bas (30% en moyenne, dont 40% pour les francophones majoritaires et 20% pour les francophones minoritaires), on peut expliquer ce phénomène par le fait que *léger* est un adjectif qui ne se place pas de manière systématique avant l'adjectif. On dirait par exemple *un repas léger*. Par contre, l'adjectif *dernier* se place, dans ce sens, toujours avant le nom, car s'il se place après, il a un autre sens. Les expressions figées *accent aigu* (taux de cooccurrence de 1 dans *Lexique*) et *fait accompli* (taux de cooccurrence de 12) sont moins fréquemment utilisées que *sous-entendu* (38) ou *États-Unis* (+ de 1000).

Le rôle de la fréquence des mots est visible à plusieurs niveaux. Les francophones de tous les groupes auront tendance à moins faire de liaisons pour les mots les moins fréquents. Par exemple, les mots pour lesquels il y a eu moins de 70% de production pour tous les francophones réunis, appartiennent tous à la catégorie K3 ou HL. Toutefois, ce n'est pas parce qu'un mot appartient à ces catégories, que les sujets ne feront pas les liaisons. Il s'agit ici d'un mélange de trois facteurs : le facteur syntaxique (la catégorie de mot ou groupes de mots), la complexité de la règle et le facteur de fréquence.

Entre un déterminant et un nom, la fréquence lexicale du nom ne semble pas faire de différence. Que ce soit *des étudiants* (K1), *mes ouvriers* (K2), ou *un accent* (K3), tous les francophones ont internalisé la règle que [dét. + nom] produit une liaison obligatoire. D'ailleurs, d'après les règles, les seuls cas où il n'y a pas de liaison ce sont avec les h aspirés et les semi-consonnes, mots pour la plupart d'origine étrangère qui constituent des exceptions à la règle. Les règles concernant les semi-consonnes étant simples (sauf pour « oiseau » et « oie » et mots de même famille, les liaisons sont interdites), tous les francophones connaissent ces exceptions même pour le mot « oie » qui est pourtant HL.

Lorsque les structures sont moins productives, alors la fréquence des mots joue un rôle plus important. Les règles concernant les liaisons dans les catégories [adj. + nom] et expressions figées sont moins productives, d'où l'effet de la fréquence : la liaison n'a été faite que par 10 francophones sur 12 avec *fait accompli*, et par 7 francophones sur 12 avec *accent aigu*.

Alors que les adjectifs possessifs, démonstratifs, interrogatifs et indéfinis sont très fréquents, certains adjectifs qualificatifs peuvent être moins fréquents. Par exemple, l'adjectif *léger* appartient à la catégorie K2. Si l'adjectif est lui-même assez rare (appartenant à une autre catégorie que K1), mais qu'en plus, le nom qu'il modifie appartient lui-même à une catégorie rare, la liaison a peu de chance de se faire. Cela explique pourquoi la liaison n'a été

faite qu'à 40% par les Européens comme pour les Québécois dans le contexte *léger ennui*.

Si un mot est peu fréquent, un francophone minoritaire peut l'avoir entendu encore moins souvent qu'un francophone majoritaire, étant donné qu'en situation de contact restreint les contextes d'apprentissage et de discussion sont évidemment moins nombreux. Autrement dit, des francophones minoritaires peuvent avoir entendu des mots appartenant à la catégorie K2 le même nombre de fois que des francophones majoritaires auront entendu des mots de la catégorie HL. Le seuil « d'écoute » et donc de production va être différent pour les francophones majoritaires et minoritaires. C'est ainsi que pour une liaison avec un mot appartenant à la catégorie K2 tel que *grand immeuble*, les francophones majoritaires ont fait la liaison dans 100% des cas alors que les Franco-ontariens ne l'ont faite que dans 25% des cas et les Africains dans 50% des cas ; de même, dans la catégorie K3 tel que *ancien édifice*, les francophones majoritaires ont fait la liaison dans 100% des cas alors que les Franco-ontariens dans 75% des cas et les Africains dans 25% des cas.

Il est d'ailleurs intéressant de remarquer que le mot *immeuble* est plutôt utilisé par les Européens et les Africains, alors que les Canadiens ont plus tendance à utiliser le mot *édifice* (même mot en anglais) à la place. Cela expliquerait peut-être pourquoi les Africains ont fait plus de liaisons que les Franco-ontariens avec *grand immeuble*, et pourquoi les Franco-ontariens ont fait plus de liaisons avec *ancien édifice* que les Africains.

D'après nos résultats, les facteurs qui jouent le rôle le plus important dans la réalisation de liaisons, sont les facteurs liés à la fréquence. La fréquence du mot1 est un indicateur du type de structure et donc de la fréquence de la structure. Si le mot1 est un article singulier, si cet article est un mot très fréquent, cela indique par la même occasion que la structure article + nom (car la plupart du temps il y a un nom après l'article) est une structure fréquente. Cela indique aussi que les articles singuliers sont fréquents. La fréquence du mot2 est un indicateur de la fréquence lexicale. Et le taux de cooccurrence indique les probabilités

de trouver ces deux mots ensemble. La fréquence de cooccurrence des deux mots est aussi un bon indicateur de la fréquence lexicale et montre plus que tout autre facteur, la fréquence à laquelle les séquences ont été entendues.

Enfin, la fréquence de la consonne de liaison joue également un rôle dans la production de liaison. Les liaisons avec des consonnes contenant un nombre trop restreint d'items finissent par disparaître (avec le g, comme dans *sang impur*), et celles qui sont très rares sont en train de disparaître (avec le r). Dans ce cas, les effets cumulés de la fréquence phonétique, syntaxique et lexicale associée au capital linguistique des locuteurs vont déterminer la production de liaison.

La productivité d'une structure joue un rôle important, notamment le nombre d'items qu'elle regroupe et le nombre d'items et de structures ayant des règles opposées à cette structure. Ainsi, puisque la structure [nom sing. + adj.] est normalement un contexte de liaison interdite (*garçon intelligent*), une expression figée comprenant ce même contexte est peu productive. De plus, les expressions figées sont en nombre restreint. La catégorie expression figée est donc peu productive, d'où un nombre plus restreint de liaisons réalisées.

Si certains facteurs tels que l'âge, le sexe, le niveau d'étude, l'origine géographique peuvent jouer un rôle dans la production de liaisons, les facteurs les plus importants, sont d'après nos résultats, les facteurs liés à la fréquence. Par exemple, les personnes qui ont obtenu un doctorat, sont les personnes qui ont probablement le plus entendu de liaisons dans leur entourage, en début d'acquisition et plus tard dans leurs études. Ce sont aussi les personnes de parents plus éduqués qui ont eux-mêmes produit plus de liaisons, etc. Il est donc difficile de séparer de tels facteurs sociolinguistiques.

Par contre, les banques de données telles que *Lexique* et *Vocabprofil* nous ont permis d'étudier les effets de la fréquence des mots et de leur cooccurrence et d'en venir à la conclusion que les facteurs les plus importants dans les productions de liaisons pour les

francophones majoritaires et minoritaires sont tous reliés à la fréquence : la fréquence syntaxique, lexicale et la fréquence des consonnes de liaison, ainsi que la productivité de ces structures, tels qu'illustré dans le graphique 4.3. ci-dessous.



**Graphique 4.3. Facteurs participant à l'acquisition des liaisons**



## 2. Les productions de liaisons et enchaînements des anglophones

Comme mentionné précédemment, d'après le MBU et d'après le niveau des étudiants de notre étude, environ le niveau intermédiaire B1, d'après le Cadre Européen Commun de Référence (CECR), notre hypothèse est qu'à ce stade de leur apprentissage, ils ne peuvent produire que quelques items individuels (lexicalisés), ceux qu'ils ont le plus souvent entendus, notamment dans le contexte le plus fréquent [art. + nom]. Nous pensons qu'ils n'ont pas entendu de séquences bien formées suffisamment de fois pour avoir acquis une base de données assez riche qui leur permette d'internaliser des règles abstraites.

Nous avons vu que les francophones se basent en grande partie sur la fréquence des items et des structures, la fréquence du mot1, la fréquence du mot2 et la fréquence de cooccurrence des mots, ainsi que sur la consonne de liaison pour déterminer quand et comment faire la liaison. Nous avons également remarqué qu'il y avait un continuum allant des francophones majoritaires (Français, Belges, Québécois) qui produisaient le plus grand nombre de liaisons, les personnes bilingues ayant un parent francophones qui en faisaient juste un peu moins, mais globalement à un niveau similaire à tous les autres francophones majoritaires, et les francophones minoritaires qui en faisaient nettement moins. Les Franco-ontariens produisaient un peu plus de liaisons que les Africains. Ce continuum semble correspondre à la fréquence d'écoute et à l'accès à la langue. Ainsi, d'après les données de notre étude sur la fréquence, nous en concluons que plus l'input des locuteurs est riche et varié, plus la fréquence d'écoute augmente et plus la production de liaisons est élevée.

Or, dans le cas d'étudiants anglophones, le nombre de liaisons entendues est probablement plus restreint que pour les francophones minoritaires. Les sujets de notre étude sont en première année d'université, aucun d'entre eux n'a poursuivi ses études dans une école francophone sauf une personne qui a fréquenté une école francophone pendant un temps limité, ce que les réponses au questionnaire ont confirmé. Les étudiants viennent du système

de français de base (*core French*), d'immersion, ou de français intensif (*Extended French*) et parfois d'un mélange de plusieurs systèmes. C'est-à-dire que certains peuvent par exemple avoir fréquenté une école d'immersion pendant toute leur scolarité, d'autres peuvent avoir étudié le français seulement en français de base, tandis que d'autres peuvent avoir étudié dans plusieurs systèmes. Aucun d'entre eux n'a commencé à étudier le français à l'université au niveau débutant.

Le français entendu dans les écoles n'est pas toujours parlé par des natifs. Le français entendu en classe vient d'autres élèves et surtout des enseignants de français souvent anglophones. Le modèle donné joue un rôle primordial, comme nous l'avons remarqué dans notre étude avec les francophones ayant une mère francophone et dans les études portant sur l'acquisition des liaisons chez les enfants francophones (Dugua, 2005, 2006). L'input reçu en début d'acquisition semble être crucial dans l'acquisition de la prononciation et des liaisons.

Nous supposons donc que si les enseignants n'étaient pas francophones, ils ont eux-mêmes produit moins de liaisons que des francophones majoritaires. Nous avons vu aussi que les francophones minoritaires produisent moins de liaisons que les francophones majoritaires. De nombreux enseignants sont des francophones minoritaires. Ainsi, si les enseignants étaient anglophones et francophones minoritaires, les étudiants auront entendu moins de liaisons que les francophones majoritaires.

De plus, les contextes d'écoute du français que les élèves ont rencontrés en salle de classe étaient plus réduits dans le nombre d'heures comme dans la variété des situations, en particulier pour les élèves en français de base. D'après ces conditions d'apprentissage et l'input forcément plus réduit que celui des francophones majoritaires, les étudiants du français L2 devraient réaliser les liaisons obligatoires, les liaisons facultatives et les enchaînements dans un nombre plus restreint que les francophones minoritaires, si l'input joue un rôle majeur.

## **2.1. Les résultats de notre étude**

Nous souhaitons ici analyser plusieurs facteurs pour mieux comprendre le processus d'acquisition des liaisons et le système de traitement des informations des apprenants du français langue seconde ou langue étrangère. Nous allons d'abord considérer les productions des anglophones « du temps 1 », ce qui correspond au niveau des apprenants avant la petite leçon de phonétique, et les comparer aux productions des francophones. Nous analyserons divers facteurs que nous comparerons aux productions, puis nous étudierons les progrès des étudiants après la leçon dans les différents groupes.

### **2.1.1. Résultats généraux**

Comme nous le voyons dans le tableau 4.29., les apprenants du français L2 ont produit un total de 52.3% de liaisons obligatoires, facultatives et interdites (liaisons correctement produites, donc non réalisées) et enchaînements. Les francophones minoritaires en ont produit 66.7% et les francophones majoritaires 81.53%. Les productions des anglophones sont nettement plus basses (52.3%) que celles des francophones majoritaires (81.53%) et un peu plus proche des francophones minoritaires (66.7%), ce à quoi nous nous attendions.

C'est avec les liaisons obligatoires que les différences sont les plus visibles entre les francophones et les apprenants. Les productions de liaisons obligatoires des francophones majoritaires de notre étude (95.6%) sont très proches des productions de francophones d'autres études telle celle de Malécot (1975) : 96.9%. Les apprenants ont produit 60.7% de liaisons obligatoires, ce qui constitue une énorme différence avec les francophones majoritaires.

De plus, les 60.7% de liaisons obligatoires des anglophones de notre étude sont également très éloignés des 91.1% de l'étude de Thomas (2002), des 95.36 % de l'étude de Mastro Monaco (2000) (ou 93.32% pour la moyenne des deux groupes pour toutes les activités) ou des 95% de l'étude de Howard (2004).

Ces différences de résultats reflètent les différences dans le contenu de ce qui a été testé. Si l'on teste uniquement les contextes les plus fréquents tels que les articles + noms, on obtient un taux de réalisation très élevé. Mais si l'on teste d'autres contextes tels que les liaisons avec différentes sortes d'adjectifs, expressions figées et autres cas pourtant obligatoires, les taux de réalisation des liaisons sont plus bas et on peut supposer qu'ils sont plus représentatifs des compétences réelles des apprenants.

Le nombre réduit de productions de liaisons des anglophones était prévisible puisque notre texte à lire comprenait une plus grande variété de contextes que les études précédentes (pour celles qui avaient un texte à lire), et plus aussi que la parole spontanée, bien qu'il s'agisse toujours de liaisons obligatoires.

Ensuite, le taux d'enchaînements est beaucoup plus bas (55.1%) que celui des francophones majoritaires (78.2%), mais légèrement plus élevé que celui des francophones minoritaires (54.8%). Les anglophones ont produit moins de liaisons facultatives (22.4%) que les francophones majoritaires (37.3%), mais un peu plus que les francophones minoritaires (19.1%).

Le nombre plus élevé de liaisons facultatives et d'enchaînements que celui des francophones minoritaires est surprenant au premier abord. Vu le nombre élevé de liaisons et enchaînements, les sujets ont dû faire des choix, et réaliser ce qu'ils jugeaient être prioritaire. Les francophones majoritaires et minoritaires réalisent en priorité les liaisons obligatoires, ensuite les enchaînements et enfin sur les liaisons facultatives, ce qu'on observe dans le taux de liaisons obligatoires très élevé pour les deux groupes (95.6% pour les francophones majoritaires et 85.5% pour les francophones minoritaires), suivi du taux d'enchaînements (78.2% pour les francophones majoritaires et 54.8% pour les francophones minoritaires) et enfin les liaisons facultatives sont nettement moins réalisées que les liaisons obligatoires (37.3% pour les francophones majoritaires et 19.1% pour les francophones minoritaires).

Les productions de liaisons (tableau 4.29.) suggèrent que les anglophones semblent savoir qu'il faut faire des liaisons, puisqu'ils les font plus de la moitié du temps (52.3% pour toutes les liaisons et enchaînements), mais ils ne semblent pas connaître les « priorités ». Ils peuvent ne pas faire une liaison obligatoire et faire juste après une liaison facultative (avec *pas* en particulier) et faire des liaisons interdites. C'est ce qui pourrait expliquer que les anglophones fassent un peu plus de liaisons facultatives et d'enchaînements et plus de liaisons interdites que les francophones minoritaires, mais nettement moins de liaisons obligatoires.

Par exemple, le sujet 1 du groupe 3, fait une liaison facultative avec *ont eu*, mais ne fait pas la liaison obligatoire avec *un accent*, alors que le contexte [dét. + nom] est pourtant « prioritaire »; les sujets 5 et 7 du groupe 3 font une liaison facultative avec *pas allés*, mais ne font pas la liaison obligatoire avec *un accent*.

Enfin, les anglophones ont réalisé correctement les liaisons interdites (ne les ont pas prononcées) dans 55.2% des cas, comparativement aux francophones majoritaires (86.3%) et aux francophones minoritaires (66.7%). C'est-à-dire qu'ils ont fait plus d'erreurs avec les liaisons interdites (en les produisant quand il ne faudrait pas) que les francophones minoritaires et beaucoup plus que les francophones majoritaires.

Paradoxalement, le fait qu'ils aient produit presque la moitié de liaisons interdites peut être considéré comme de « bonnes erreurs », représentatives d'une mise en place d'un système. S'ils ont fait des liaisons interdites, cela signifie que dans leur lexique, la règle générale est de faire des liaisons dans ces contextes et qu'ils ont une certaine connaissance de la contrainte anti-hiatus et des règles générales de liaisons. Seulement ils ne connaissent pas tous les cas particuliers (si difficiles à connaître, même pour les francophones), qui se trouvent d'ailleurs souvent avec des mots relativement rares.

	francophones						anglophones								
	MAJ			MIN			T 1			T 2			T 3		
tous	L	N	%	L	N	%	L	N	%	L	N	%	L	N	%
enchaîn.	122	156	<b>78,2</b>	57	104	<b>54,8</b>	265	481	<b>55,1</b>	305	481	<b>63,4</b>	293	481	<b>60,9</b>
L. oblig.	585	612	<b>95,6</b>	349	408	<b>85,5</b>	1145	1887	<b>60,7</b>	1304	1887	<b>69,1</b>	1311	1887	<b>69,5</b>
L. fac.	76	204	<b>37,3</b>	26	136	<b>19,1</b>	141	629	<b>22,4</b>	161	629	<b>25,6</b>	167	629	<b>26,6</b>
L. int.	145	168	<b>86,3</b>	75	112	<b>67,0</b>	286	518	<b>55,2</b>	295	518	<b>56,9</b>	289	518	<b>55,8</b>
<b>total</b>	<b>928</b>	<b>1140</b>	<b>81,55</b>	<b>507</b>	<b>760</b>	<b>66,70</b>	<b>1837</b>	<b>3515</b>	<b>52,30</b>	<b>2065</b>	<b>3515</b>	<b>58,70</b>	<b>2060</b>	<b>3515</b>	<b>58,60</b>

Tableau 4.29. Liaisons obligatoires, facultatives, interdites et enchaînements francophones et anglophones

Ce qui distingue le plus les apprenants des francophones, c'est la production de liaisons obligatoires et de liaisons interdites, car c'est dans ces domaines que la norme orthoépique est la plus précise et qu'on peut véritablement parler d'erreurs. Les enchaînements et les liaisons facultatives sont également produits en plus faible quantité que chez les francophones majoritaires.

Étant donné que les anglophones ont un taux d'enchaînements plus bas que les francophones, l'une des questions qui se posent est la suivante : les anglophones font-ils moins de liaisons (et produisent-ils moins de liaisons enchaînées), à cause du système syllabique différent, et à cause de l'interférence de l'anglais en général? Pour répondre à cette question, nous avons considéré le nombre de liaisons non-enchaînées et les productions globales selon les différentes langues maternelles des apprenants.

## 2.1.2. Les liaisons selon les groupes linguistiques

### 2.1.2.1. Les liaisons non-enchaînées

Une grande différence entre les productions des francophones et des anglophones est les liaisons non-enchaînées (tableaux 4.30. et 4.31.). Les anglophones font aussi moins d'enchaînements que les francophones majoritaires, soulignant la possibilité de l'influence du système syllabique de leur L1 (l'anglais) sur le français.

Les anglophones produisent 8.85% de liaisons obligatoires non-enchaînées (avant la

petite leçon), soit 167 liaisons non-enchaînées sur 1887 occurrences alors que les francophones majoritaires ne produisent que 2 liaisons enchaînées sur 1020 occurrences soit 0.2%. Autrement dit, les anglophones produisent 44.25 fois plus de liaisons non-enchaînées que les francophones majoritaires.

Chez les francophones, les liaisons non-enchaînées semblent dues aux hésitations et sont plutôt des erreurs de performance au même titre que les reprises, comme l'a observé Mallet (2008), contrairement aux liaisons non-enchaînées des hommes politiques étudiés par Encrevé (1988) qui faisaient exprès de ne pas enchaîner les liaisons pour des questions de style et de prestige. Ce type de prononciation consistant à ne pas enchaîner les liaisons est réservé à une élite (les hommes politiques et les personnages médiatiques) et ne s'applique pas à la prononciation du francophone « moyen » comme en témoignent par exemple les résultats du projet PFC.

Le fait que les liaisons non-enchaînées soient un signe d'hésitation chez les francophones se confirme par un taux plus élevé de liaisons non-enchaînées avec les liaisons facultatives (0.59%), soit deux fois plus (proportionnellement) qu'avec les liaisons obligatoires. Le taux de fréquence des structures et des items des liaisons facultatives est plus bas que celui des liaisons obligatoires, ce qui crée plus d'hésitations qu'avec les liaisons obligatoires, et ainsi plus de liaisons non-enchaînées.

Chez les anglophones, les liaisons non-enchaînées sont dues à d'autres raisons que les reprises et hésitations. Nous pensons que plusieurs facteurs participent à ce phénomène. Tout d'abord le système de syllabation est différent en français et en anglais, ce qui crée des problèmes d'interférence dans la production des anglophones. Delattre (1946) avait déjà remarqué ce phénomène :

Mais l'opposition entre l'anglais et le français est ici à la fois si profonde et si subtile que pour acquérir des habitudes de division syllabique française il faudra longtemps en exagérer l'articulation. Les syllabes françaises sont presque invariablement du type appelé "ouvertes," type qui doit son nom au fait que la bouche est littéralement ouverte quand la syllabe se termine, et non fermée comme en anglais. (Delattre, 1946 : 45)

Le fait que les anglophones aient des difficultés à enchaîner les consonnes en français (que ce soit pour les enchaînements comme pour les liaisons enchaînées) est un phénomène bien connu des phonéticiens. D'où par exemple les conseils de Duménil (2003 :10) pour les anglophones : « [...] les locuteurs français anticipent les voyelles, tandis que les locuteurs de langue anglaise anticipent les consonnes. L'apprenant américain devra donc redoubler d'efforts pour ne pas anticiper les consonnes dans sa prononciation du français [...] afin de maximiser le nombre de syllabes ouvertes dans sa prononciation. »

Toutefois les différences de syllabation du français et de l'anglais ne semblent pas être la seule cause à ce nombre élevé de consonnes finales prononcées et non-enchaînées. Si tel était le cas, on trouverait soit un pourcentage similaire de liaisons obligatoires et de liaisons facultatives non-enchaînées, soit un nombre plus élevé de liaisons facultatives non-enchaînées. Les liaisons facultatives se trouvant dans des contextes moins fréquents devraient être moins familières pour les anglophones, qui auraient alors recours à l'utilisation du système syllabique de leur langue maternelle, l'anglais. On trouverait alors toutes les consonnes prononcées, ou au moins certaines consonnes prononcées de façon systématique en particulier dans les contextes plus rares.

Or, ce n'est pas le cas. Le pourcentage de liaisons obligatoires non-enchaînées est beaucoup plus élevé (8.85%) que le pourcentage de liaisons facultatives non-enchaînées (1.43%). Cela suggère que le non-enchaînement des liaisons n'est pas un phénomène général applicable à toutes les consonnes finales, indépendamment de leur statut (en liaison obligatoire ou facultative) et de leur fréquence.

De plus, d'après ce que nous avons pu observer des enregistrements de notre étude, les anglophones ont tendance, dans certains cas, à prononcer les consonnes finales, comme ils le feraient en anglais, y compris lorsque le mot est suivi d'une consonne. Par exemple, ils prononcent ils vont [ilz vɔ̃]. Ils ne semblent pas connaître de manière formelle les règles



phonétiques qui demandent soit de ne pas prononcer les consonnes finales, soit de les prononcer enchaînées avec la liaison. S'ils connaissaient ces règles, ils ne feraient pas un aussi grand nombre de liaisons non-enchaînées qui sont en fait dues à la prononciation de la consonne finale, indépendamment de la liaison.

Nous avons noté en particulier, un grand nombre de /n/ prononcé et non-enchaîné dans l'article *un*, ce qui est une sorte d'article « épïcène » se situant entre le masculin et le féminin, qui pourrait être interprété comme étant le calque de l'anglais *an*. Par exemple, le /n/ de *un* est prononcé et non-enchaîné (tableau 4.30) 11 fois dans *un accent*, 7 fois dans *un hôpital*, 1 fois dans *un oiseau*. Lightbown et Spada (2006 : 173) mentionnent ce phénomène de l'article hybride : « [...] because previous research had shown that learners sometimes used a 'hybrid article' that could be interpreted as either masculine or feminine. »

En outre, le /n/ est problématique dans d'autres cas que dans l'article *un*. Le /n/ de *certain âge* est prononcé et non-enchaîné 8 fois, le /n/ de *ancien édifice*, 6 fois, de *aucun effort*, 5 fois, et de *son école*, 4 fois, ce qui semble indiquer que les nasales sont problématiques en général, que ce soit avec *un* ou dans d'autres cas. D'ailleurs, la prononciation de consonne finale se remarque avec d'autres phonèmes, ce qui souligne un problème phonétique général.

La lecture présente alors un autre avantage, par rapport à la parole spontanée. Il n'y a, en lecture, aucune ambiguïté de genre (*un* vs *une*) puisque l'article détermine le genre du nom. Si dans la parole spontanée on pouvait interpréter cette prononciation comme provenant de fautes grammaticales, la lecture enlève toute autre interprétation quant aux causes de cette prononciation. Les causes d'erreurs sont clairement phonétiques.

Un autre facteur qui pourrait influencer le non-enchaînement des consonnes de liaison serait le mélange de plusieurs facteurs phonétiques. Si les apprenants accentuent la première syllabe du mot<sup>2</sup> (au lieu de la dernière syllabe du dernier mot du groupe), comme cela serait

le cas avec l'accent d'insistance chez les francophones, cet accent tonique pourrait inciter le locuteur à produire un coup de glotte qui contribuerait au non-enchaînement de la consonne.

Ce système différent d'accentuation de syllabe a déjà été remarqué par Delattre (1946) :

L'accentuation de syllabes en anglais aura pour effet que les Anglophones auront tendance à inégaliser les syllabes et utiliser ce même accent en français. En français, il y a l'accent final qui n'est pas du tout comparable à l'accent anglais ni dans sa nature ni dans sa place. L'accent d'insistance est plus proche de l'accent anglais mais il est employé très rarement. De ce fait, il pourra y avoir des difficultés d'accentuation de syllabes. (Delattre, 1946 : 48)

Si les apprenants prononcent les consonnes de liaison de façon similaire à l'anglais, cela pourrait aussi empêcher l'enchaînement. Par exemple, il est difficile d'enchaîner le /R/ de *léger* avec *ennui*, si le /R/ est prononcé comme en anglais. Pour déterminer si la langue maternelle et les différences de prononciation des consonnes de la L1 et de la L2 peuvent jouer un rôle dans la production de liaisons, nous avons analysé un peu plus loin les productions de consonnes de liaison et les avons comparées à différents groupes ayant une L1 différente.

Enfin, si les mots sont inconnus ou peu connus, que les apprenants n'ont pas l'habitude de lire à haute voix, une tendance de tous les apprenants est de prononcer les mots séparément, et ce, indépendamment de la L1 et de la L2. On voit cela par exemple dans l'étude de Hieke (1984) avec les apprenants germanophones de l'anglais.

L'interférence de la langue maternelle pourrait être un facteur dans la prononciation des consonnes finales, dans le non-enchaînement des liaisons et dans le nombre restreint de liaisons et d'enchaînements, mais il semble aussi y avoir un autre facteur pour expliquer cela. De façon inattendue, le pourcentage de liaisons obligatoires non-enchaînées baisse après la leçon de phonétique (tableau 4.30.), alors que le pourcentage de liaisons facultatives non-enchaînées augmente légèrement après la leçon (tableau 4.31.).

Rappelons que lors de la leçon de phonétique, les étudiants ont reçu des explications rapides sur les liaisons et les enchaînements. La leçon portait principalement sur les liaisons

obligatoires et aussi sur les cas les plus fréquents de liaisons interdites. Il a été mentionné qu'on ne doit pas prononcer la consonne finale, sauf en liaison. Plusieurs exemples de liaisons enchaînées et d'enchaînements ont été donnés. Il y a aussi eu une brève explication de ce qu'il ne faut pas faire, avec des exemples : des consonnes finales prononcées quand il ne faudrait pas, avec mise en évidence avec des gestes, des explications et avoir écrit au tableau et barré la consonne finale, pour montrer qu'il ne faut pas prononcer la consonne finale si elle n'est pas en contexte de liaison. Les règles sont simples, faciles à comprendre.

En effet, à la question du questionnaire: comment les concepts utilisés pour cette présentation vous ont-ils semblé ? Les réponses possibles étaient : très difficiles à comprendre, assez difficiles à comprendre, difficiles à comprendre, faciles à comprendre, très faciles à comprendre. 4 sujets ont répondu « très faciles », 31 ont répondu « faciles », et 2 ont répondu « difficiles ». En d'autres termes, 97.59% des apprenants ont trouvé la présentation facile ou très facile à comprendre. Les concepts expliqués étaient simples à comprendre : on ne prononce pas la consonne finale s'il n'y a pas de *e* (schwa), à moins qu'on la prononce en liaison AVEC le mot d'après. Il est peu probable que les apprenants aient confondu ou mal compris les règles.

LO	nb LNE (c)	total	%
tous franc	2	1020	<b>0,20</b>
ang temps1	167	1887	<b>8,85</b>
ang temps2	122	1887	<b>6,47</b>
ang temps3	126	1887	<b>6,68</b>

**Tableau 4.30. Liaisons obligatoires non-enchaînées, francophones et anglophones**

LF	nb LNE(c)	total	%
tous franc	2	340	<b>0,59</b>
ang temps1	9	629	<b>1,43</b>
ang temps2	12	629	<b>1,91</b>
ang temps3	11	629	<b>1,75</b>

**Tableau 4.31. Liaisons facultatives non-enchaînées, francophones et anglophones**

Si l'on observe le tableau 4.32., dans lequel les liaisons non-enchaînées ont été classées par ordre décroissant, on voit que la consonne de liaison ne semble pas déterminer à elle seule le non-enchaînement des consonnes, puisqu'on retrouve ce phénomène avec tous les phonèmes. 7 /v/, 22 /R/, 22 /t/, 49 /n/ et 67 /z/ n'ont pas été enchaînés. Le /z/ ou le /v/ ne présentent pas de difficultés articulatoires particulières pour les sujets (contrairement au /R/).

Les seules séquences qui ont toutes été enchaînées par tous les sujets (si elles ont été prononcées) sont *nos amis*, *vieux écrits*, *aux yeux*, *deux amis*, *trois ans*. Pourquoi les sujets font-ils un enchaînement systématique avec deux et trois, mais pas avec dix ? Il s'agit pourtant du même contexte syntaxique et du même contexte phonétique (phonème /z/).

Il est possible que le non-enchaînement des consonnes de liaison soit relié à l'effet de fréquence (syntaxique et lexicale). Par exemple, la liaison des adjectifs qualificatifs + noms n'est réalisée qu'à 86.6% par les locuteurs francophones, et cette structure est une structure moins fréquente que les autres structures des liaisons obligatoires. Quant à la fréquence lexicale, la séquence *léger ennui* avait un taux de cooccurrence faible et les deux mots étaient moyennement fréquents. Cela a été la séquence de liaison obligatoire la moins réalisée par les francophones (30%) est précisément la séquence ayant engendré le plus de liaisons non-enchaînées de la part des étudiants L2. La fréquence pourrait donc jouer un rôle dans la prononciation de consonnes finales et de consonnes non-enchaînées chez les anglophones apprenant le français.

Toutefois, la fréquence à elle seule ne permet pas d'expliquer pourquoi les sujets n'ont pas enchaîné les liaisons dans des séquences aux fréquences syntaxique et lexicale élevées telles que *des histoires*, *dans un*, *six heures*. Il est possible que la productivité des structures rentre aussi en jeu.

	liaisons non enchaînées	g1	g2	g3	g4	total
LO18	léger ennui	4	4	3	1	12
LO34	un accent	1	3	6	1	11
LO10	certain âge	4	3	1	0	8
LO37	un hôpital	1	4	1	1	7
LO23	grand immeuble	1	5	0	1	7
LO49	neuf heures	0	0	7	0	7
LO16	premier amour	2	2	2	0	6
LO20	ancien édifice	3	2	0	1	6
LO29	les oies	3	1	1	1	6
LO14	aucun effort	1	3	1	0	5
LO4	Tes assistants	3	1	1	0	5
LO43	dix ans	3	1	1	0	5
LO19	dernier étage	2	2	0	0	4
LO15	son école	0	1	3	0	4
LO48	vingt cinq	0	0	4	0	4
LO26	belles acclamations	2	1	1	0	4
LO40	des histoires	1	0	2	1	4
LO41	ces hôpitaux	2	1	1	0	4
LO51	dans un	2	1	1	0	4
LO17	fait accompli	0	0	3	0	3
LO22	second étage	1	1	1	0	3
LO1	aux invités	3	0	0	0	3
LO6	leurs enfants	1	2	0	0	3
LO25	nouvelles idées	0	0	3	0	3
LO39	vieilles éponges	1	1	1	0	3
LO44	six heures	0	0	3	0	3
LO45	trois arbres	2	0	1	0	3
LO9	moyen-âge	2	0	0	0	2
LO12	mon avis	0	0	2	0	2
LO35	accent aigu	0	0	2	0	2
LO46	vingt trois	0	0	2	0	2
LO3	des étudiants	1	1	0	0	2
LO5	mes ouvriers	2	0	0	0	2
LO21	nouveaux étudiants	1	0	1	0	2
LO31	folles aventures	0	1	1	0	2
LO33	sous-entendu	1	1	0	0	2
LO36	États-Unis	0	1	1	0	2
LO38	vieux habits	1	0	1	0	2
LO8	bon ami	0	0	1	0	1
LO11	plein air	1	0	0	0	1
LO13	ton enfant	0	0	1	0	1
LO27	un oiseau	0	0	1	0	1
LO50	cent ans	0	0	1	0	1
LO2	les affaires	0	1	0	0	1
LO30	beaux Allemands	0	1	0	0	1
LO32	faux-airs	0	0	0	1	1
LO7	nos amis	0	0	0	0	0
LO24	vieux écrits	0	0	0	0	0
LO28	aux yeux	0	0	0	0	0
		52	45	62	8	167

**Tableau 4.32. Liaisons non-enchaînées par occurrences et groupes**

Nous avons vu dans les chapitres 1 et 2 que la productivité et la complexité d'une structure jouent un rôle majeur dans l'acquisition. Par exemple, si une structure 1 regroupe plus d'items qu'une structure 2, la structure 1 rentre en compétition avec la structure 2 qui a moins d'items. L'enfant va d'abord apprendre la structure 1 qui regroupe plus d'items, surgénéraliser cette règle, et l'utiliser là où il faudrait utiliser la structure 2. D'autre part, si une structure a des exceptions, et des règles différentes selon les cas, la structure est complexe.

Par exemple la structure grapho-phonémique de *dix* est peu productive, parce qu'il y a peu de chiffres se terminant par x (*six, dix*) et comprend donc un nombre d'items très restreint, et la structure est complexe dans la mesure où le x peut être prononcé /s/, /z/ ou ne pas être prononcé du tout (dans *dix livres*), alors que dans d'autres cas d'ailleurs restreints (aux), x en liaison est toujours prononcé /z/. Il en est de même avec neuf heures. D'une part, il n'existe qu'un seul cas où le f se prononce /v/, c'est avec *neuf*. C'est donc une structure peu productive. D'autre part, le f de *neuf* ne se prononce /v/ que dans deux cas : *neuf ans* et *neuf heures*. Dans tous les autres cas, le f se prononce /f/ comme dans *neuf amis, neuf anges*. La structure est donc complexe et peu productive.

La comparaison des anglophones aux autres groupes linguistiques permettra de nous éclairer sur le rôle de la langue maternelle dans la production de liaisons et enchaînements.

### **2.1.2.2. Résultats généraux par groupes linguistiques**

Après avoir observé les liaisons non-enchaînées de tous les apprenants comparées aux francophones, voyons maintenant plus précisément la relation entre la langue maternelle, le non-enchaînement des liaisons et sur les productions générales.

Les locuteurs de notre étude dont la langue maternelle est autre que l'anglais ou le chinois et qui n'ont qu'une seule langue maternelle (l'allemand ou l'arabe) produisent le plus

grand nombre de liaisons non-enchaînées (6.71%) (tableau 4.33.) et paradoxalement, ce sont eux qui produisent quand même le plus grand nombre de liaisons et enchaînements (tableau 4.37.) et de liaisons obligatoires (tableau 4.38.) Leur production de liaisons obligatoires enchaînées atteint 68.6% au temps 1, 86.3% au temps 2 et 82.4% au temps 3, ce qui se rapproche de la production de liaisons obligatoires des francophones minoritaires (85.5%).

Le locuteur chinois (dont le chinois est la seule langue maternelle) n'a produit aucune liaison obligatoire non-enchaînée (tableau 4.34.), et aucune liaison facultative non-enchaînée (tableau 4.35.) au premier enregistrement. Il a produit un total de deux liaisons enchaînées (2 sur 246 contextes en prenant en compte les 82 liaisons multipliées par les trois enregistrements). Il a pourtant produit 74.5% de liaisons obligatoires ce qui est un pourcentage plus bas que les locuteurs allemands et arabes (82.5%) et très proche des locuteurs bilingues (76.8%) (tableau 4.38.). Le non-enchaînement des consonnes n'est donc pas le principal facteur qui empêche la production « correcte » de liaisons, et il n'est pas non plus un signe global de non maîtrise des liaisons.

Si c'était le cas, le groupe qui a produit le plus grand nombre de liaisons non-enchaînées aurait le taux le plus bas de liaisons correctement réalisées, alors que c'est précisément le contraire. Ce groupe semble maîtriser les liaisons obligatoires à un très haut niveau, proche de celui des francophones minoritaires, et fait pourtant un grand nombre de liaisons non-enchaînées.

Les sujets bilingues<sup>42</sup> (parlant en plus de l'anglais, le créole, le cantonais, le kashi, le mandarin, le punjabi, le serbe ou le suédois) ont fait beaucoup moins de liaisons obligatoires non-enchaînées (5.88%) que les anglophones monolingues (9.80%) (tableau 4.34.). Par contre les sujets parlant l'allemand et l'arabe produisent plus de liaisons obligatoires non-enchaînées (10.78%) que les anglophones (9.80%) (tableau 4.34.). D'après Delattre (1965 : 42),

---

<sup>42</sup> Sujets qui à la question Quelle est votre langue maternelle ? ont répondu anglais + une autre langue. C'étaient des personnes nées au Canada de parents parlant une autre langue.

l'allemand a un pourcentage de syllabes ouvertes plus bas (37%) que l'anglais (40%) qui en a lui-même un pourcentage beaucoup plus faible que le français (76%).

La langue maternelle pourrait donc être un facteur dans certains aspects de la syllabation du français langue seconde ou langue étrangère, notamment le non-enchaînement des consonnes de liaison. Toutefois la L1 ne permet pas d'expliquer à elle seule pourquoi au final, ce sont les anglophones qui font le plus de liaisons non-enchaînées (7.41%), alors que les locuteurs de l'allemand et de l'arabe qui au départ en faisaient plus que les anglophones (10.78%), en font moins après la leçon de prononciation (6.86%).

De plus, les locuteurs bilingues qui ne faisaient aucune liaison facultative non-enchaînée la première fois, en ont fait trois à la deuxième étape, et les locuteurs allemands et arabes qui n'en faisaient pas au début en ont fait deux à la deuxième étape pour les liaisons facultatives alors qu'ils en ont fait moins pour les liaisons obligatoires.

Le fait que certains sujets fassent un peu plus de liaisons facultatives et de liaisons interdites non-enchaînées (tableaux 4.33., 4.34, 4.35.) après la leçon, alors qu'ils en font moins avec les liaisons obligatoires suggère qu'il s'agit peut-être d'une étape dans le développement de l'interlangue des apprenants. Ceux-ci semblent prononcer les consonnes finales avec les mots plus fréquents et les plus familiers. Après s'être familiarisés avec les mots du texte, l'ayant lu une première fois la semaine précédant le deuxième enregistrement, ces mots leur deviennent plus familiers, entraînant ainsi la prononciation de la consonne finale des liaisons facultatives qui n'était pas prononcée auparavant, alors que certaines liaisons obligatoires deviennent familières et sont maintenant enchaînées.

Le scénario développemental des apprenants serait le suivant : au stade 1 ils ne prononcent pas les consonnes finales des mots inconnus, au stade 2 ils prononcent la consonne finale non-enchaînée des mots moyennement connus et au stade 3 ils prononcent la consonne finale enchaînée avec les séquences souvent entendues, s'ils arrivent à ce dernier



stade. Il est très fréquent qu'une fossilisation s'opère au stade 2 et que les apprenants n'atteignent jamais le stade 3.

Si l'on considère les productions globales (toutes liaisons et tous enchaînements) selon les L1 des sujets (tableaux 4.33., 4.34., 4.35., 4.36.), la langue maternelle semble jouer un certain rôle puisque les locuteurs bilingues, le locuteur chinois et les locuteurs d'autres langues ont fait plus de progrès dans leur production de liaisons que les anglophones monolingues (58.7%) : 65.71% pour les bilingues, 66.32% pour les locuteurs « autres » et 68.42% pour le locuteur chinois, comparé à 66.7% pour les francophones minoritaires.

Si l'on compare les productions de liaisons et enchaînements (tableaux 4.37) des différents groupes linguistiques, ce sont les locuteurs d'autres langues que l'anglais qui en ont produit le plus grand nombre (59.47%), suivi des locuteurs bilingues (56.84%) et du sujet chinois (56.84%). Ce sont les sujets anglophones qui en ont produit le moins (50.37%). Cette tendance se manifeste d'autant plus pour les liaisons obligatoires : les locuteurs d'autres langues (que l'anglais et qu'une langue asiatique) en ont produit 68.6%, les locuteurs bilingues, 66.1%, le locuteur chinois 60.8% et enfin les anglophones 58.7%.

Toutes liaisons	T 1			T 2			T 3		
	nb	total	%	Nb	total	%	nb	total	%
anglais	142	2214	<b>6.41</b>	188	2214	<b>8.49</b>	140	2214	<b>6.32</b>
bilingue : anglais +	24	574	<b>4.18</b>	23	574	<b>4.01</b>	22	574	<b>3.83</b>
chinois	1	82	<b>1.22</b>	1	82	<b>1.22</b>	0	82	<b>0.00</b>
autre langue	11	164	<b>6.71</b>	15	164	<b>9.15</b>	13	164	<b>7.93</b>
total	178	3034	<b>5.87</b>	227	3034	<b>7.48</b>	175	3034	<b>5.77</b>

Tableau 4.33. Liaisons obligatoires, facultatives et interdites non-enchaînées par groupe linguistique

L. obligatoires	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
anglais	135	1377	<b>9.80</b>	102	1377	<b>7.41</b>	102	1377	<b>7.41</b>
bilingue : anglais +	21	357	<b>5.88</b>	14	357	<b>3.92</b>	17	357	<b>4.76</b>
chinois	0	51	<b>0.00</b>	1	51	<b>1.96</b>	0	51	<b>0.00</b>
autre langue	11	102	<b>10.78</b>	5	102	<b>4.90</b>	7	102	<b>6.86</b>
total	167	1887	<b>8.85</b>	122	1887	<b>6.47</b>	126	1887	<b>6.68</b>

Tableau 4.34. Liaisons obligatoires non-enchaînées par groupe linguistique

L. facultatives	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
liaisons non-ench c									
anglais	9	459	<b>1.96</b>	7	459	<b>1.53</b>	8	459	<b>1.74</b>
bilingue : anglais +	0	119	<b>0.00</b>	3	119	<b>2.52</b>	2	119	<b>1.68</b>
chinois	0	17	<b>0.00</b>	0	17	<b>0.00</b>	0	17	<b>0.00</b>
autre langue	0	34	<b>0.00</b>	2	34	<b>5.88</b>	1	34	<b>2.94</b>
total	9	629	<b>1.43</b>	12	629	<b>1.91</b>	11	629	<b>1.75</b>

**Tableau 4.35. Liaisons facultatives non-enchaînées par groupe linguistique**

L. interdites	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
liaisons non-ench c									
anglais	44	378	<b>11.64</b>	33	378	<b>8.73</b>	30	378	<b>7.94</b>
bilingue : anglais +	2	98	<b>2.04</b>	7	98	<b>7.14</b>	3	98	<b>3.06</b>
chinois	1	14	<b>7.14</b>	0	14	<b>0.00</b>	0	14	<b>0.00</b>
autre langue	4	28	<b>14.29</b>	4	28	<b>14.29</b>	5	28	<b>17.86</b>
total	51	518	<b>9.85</b>	44	518	<b>8.49</b>	38	518	<b>7.34</b>

**Tableau 4.36. Liaisons interdites non-enchaînées par groupe linguistique**

Toutes liaisons + ench	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
Langue maternelle									
anglais	1292	2565	<b>50.37</b>	1411	2565	<b>55.01</b>	1432	2565	<b>55.83</b>
bilingue : anglais +	378	665	<b>56.84</b>	458	665	<b>68.87</b>	437	665	<b>65.71</b>
chinois	54	95	<b>56.84</b>	57	95	<b>60.00</b>	65	95	<b>68.42</b>
autre langue	113	190	<b>59.47</b>	139	190	<b>73.16</b>	126	190	<b>66.32</b>
total	1837	3515	<b>52.26</b>	2065	3515	<b>58.75</b>	2060	3515	<b>58.61</b>

**Tableau 4.37. Productions de liaisons et enchaînements par langue maternelle**

L. obligatoires	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
Langue maternelle									
anglais	808	1377	<b>58.7</b>	899	1377	<b>65.3</b>	915	1377	<b>66.4</b>
bilingue : anglais +	236	357	<b>66.1</b>	284	357	<b>79.6</b>	274	357	<b>76.8</b>
chinois	31	51	<b>60.8</b>	33	51	<b>64.7</b>	38	51	<b>74.5</b>
autre langue	70	102	<b>68.6</b>	88	102	<b>86.3</b>	84	102	<b>82.4</b>
total	1145	1887	<b>60.7</b>	1304	1887	<b>69.1</b>	1311	1887	<b>69.5</b>

**Tableau 4.38. Productions de liaisons obligatoires par langue maternelle**

L. facultatives	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
Langue maternelle									
anglais	97	459	<b>21.1</b>	104	459	<b>22.7</b>	107	459	<b>23.3</b>
bilingue : anglais +	29	119	<b>24.4</b>	42	119	<b>35.3</b>	45	119	<b>37.8</b>
chinois	3	17	<b>17.6</b>	3	17	<b>17.6</b>	6	17	<b>35.3</b>
autre langue	12	34	<b>35.3</b>	12	34	<b>35.3</b>	9	34	<b>26.5</b>
total	141	629	<b>22.4</b>	161	629	<b>25.6</b>	167	629	<b>26.6</b>

**Tableau 4.39. Productions de liaisons facultatives par langue maternelle**

L. interdites	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
Langue maternelle									
anglais	205	378	<b>54.2</b>	104	378	<b>27.5</b>	205	378	<b>54.2</b>
bilingue : anglais +	57	98	<b>58.2</b>	42	98	<b>42.9</b>	61	98	<b>62.2</b>
chinois	8	14	<b>57.1</b>	3	14	<b>21.4</b>	9	14	<b>64.3</b>
autre langue	16	28	<b>57.1</b>	12	28	<b>42.9</b>	20	28	<b>71.4</b>
total	286	518	<b>55.2</b>	161	518	<b>31.1</b>	295	518	<b>56.9</b>

**Tableau 4.40. Productions de liaisons interdites par langue maternelle**

enchaînements	T 1			T 2			T 3		
	nb	total	%	nb	total	%	nb	total	%
Langue maternelle									
anglais	182	351	<b>51.9</b>	203	351	<b>57.8</b>	206	351	<b>58.7</b>
bilingue : anglais +	56	91	<b>61.5</b>	71	91	<b>78.0</b>	61	91	<b>67.0</b>
chinois	12	13	<b>92.3</b>	12	13	<b>92.3</b>	12	13	<b>92.3</b>
autre langue	15	26	<b>57.7</b>	19	26	<b>73.1</b>	14	26	<b>53.8</b>
total	265	481	<b>55.1</b>	305	481	<b>63.4</b>	293	481	<b>60.9</b>

**Tableau 4.41. Productions d'enchaînements par langue maternelle**

On ne peut toutefois pas affirmer que la langue maternelle est la seule façon d'expliquer pourquoi les personnes issues de diverses origines produisent plus de liaisons que les anglophones parce que plusieurs facteurs dépendent les uns des autres. Il est possible que les personnes ayant étudié dans d'autres pays aient eu un enseignement plus axé sur la forme depuis le début (pour plus de renseignements sur ce type d'enseignement, consulter Lightbown and Spada (2006 : 138-141)).

Cela pourrait expliquer qu'ils aient fait plus de progrès que les anglophones, s'ils sont habitués à faire attention à la forme. Mais ceci n'expliquerait pas pourquoi les bilingues nés au Canada et ayant été formés dans le même système que les anglophones monolingues (ayant appris le français avec l'approche la plus souvent utilisée au Canada, l'approche communicative) ont fait plus de progrès que les anglophones monolingues, alors qu'ils ont forcément étudié avec les mêmes approches pédagogiques et dans un contexte similaire.

Il se peut que cela soit dû au fait que les apprenants bilingues de naissance<sup>43</sup> ou ceux qui ont appris l'anglais L2 plus tard (les locuteurs du cantonais, de l'arabe et de l'allemand)

<sup>43</sup> Bilingues de naissance : personnes ayant appris les deux langues en même temps, dès la naissance, ou dans la prime enfance

parlent au moins trois langues : leur langue maternelle, l'anglais et le français. Ainsi, parler deux langues faciliterait l'apprentissage d'une troisième langue, notamment au niveau de la prononciation (voir Cenoz, 2003 et Hammarberg, 2009 pour un résumé des études sur l'influence interlinguistique des L1 et L2 sur une L3).

Toutes les études ne concluent pas forcément que parler deux langues aide à en apprendre une troisième. Les résultats des études sur le trilinguisme dépendent en partie du type de recherche et de ce qui a été analysé : la communication globale, les compétences globales, ou la maîtrise d'un point particulier de la troisième langue. Toutefois, la plupart des chercheurs s'accordent à dire que les sujets bilingues ont de meilleurs résultats dans l'apprentissage d'une L3 que les sujets monolingues pour apprendre cette même langue, qui est alors pour eux une L2.

En ce qui concerne le français, Bild et Swain (1989) ont comparé les résultats globaux de trois groupes : 16 enfants anglophones monolingues, 16 enfants bilingues (anglais L1, italien L2) et 15 enfants bilingues parlant une langue non romane comme L1 et l'anglais comme L2. Les deux groupes d'enfants bilingues ont eu de meilleurs résultats que les enfants monolingues. Il n'y avait pas de différences significatives entre les deux groupes bilingues, bien qu'une des langues parlées (l'italien) soit plus proche du français.

En ce qui concerne les tests de discrimination auditive, des études telles que celles de Enomoto (1994) ont également souligné les meilleurs résultats des bilingues, même si les langues parlées sont éloignées de la L3. Cenoz (2003) conclut son article en expliquant pourquoi les bilingues ont de meilleurs résultats :

Most studies on general aspects of proficiency indicate that bilingualism has a positive effect on third language acquisition when L3 acquisition takes place in additive contexts and bilinguals have acquired literacy skills in both their languages. This effect can be explained as related to learning strategies, metalinguistic awareness and communicative ability, but it can also be linked to the fact that bilinguals have a wider linguistic repertoire that can be used as a basis in third language acquisition. This positive transfer between languages is enhanced if languages are typologically close [...]. (Cenoz, 2003 : 83)

Nous avons vu avec le MBU, que les apprenants basent leur expérience et leurs compétences sur une banque de données. On peut supposer que si un locuteur a deux bases de données, il a plus de mots, de règles de grammaire, de règles phonétiques et de phonèmes dans sa banque de données qu'un sujet monolingue. Ainsi, même si des sujets ne sont pas complètement bilingues, avoir appris une L2 pourrait les aider à apprendre certains aspects d'une L3, tels que les liaisons.

### **2.1.2.3. Résultats par consonnes de liaison**

Pour déterminer si la langue maternelle (et plus précisément les phonèmes présents dans la L1 et la L2) joue un rôle dans la prononciation de consonnes finales et du non-enchaînement des consonnes, nous avons comparé les productions des sujets selon les consonnes de liaison et selon leur(s) langue(s) maternelle(s). Si la langue maternelle joue un rôle dans le traitement de la liaison, cela devrait se manifester dans le traitement des consonnes de liaison. Les différents groupes linguistiques devraient traiter certains phonèmes de la langue cible différemment selon qu'ils sont présents dans leur L1, ou proches des phonèmes de leur L1.

L'un des phonèmes du français les plus difficiles à prononcer pour les anglophones est le /R/ car il est très différent du /r/ anglais. Si l'on considère les productions des différents groupes, on observe une grande différence entre le groupe anglophone qui produit 41.98% de liaisons avec le /R/, le groupe bilingue qui en produit 47.62%, le locuteur chinois qui en produit 66.67% et le groupe « autre » qui en produit 83.3%.

Notons que le groupe parlant d'autres langues qui a produit le plus de liaisons avec le /R/ est celui pour qui le phonème est le plus proche de sa langue maternelle. Le /R/ français est plus proche du *ch* allemand et du ħ (ghain) arabe. Ceci pourrait expliquer pourquoi ce

groupe (allemand et arabe) a produit 83.33% de liaisons obligatoires avec le /R/, le locuteur chinois 66.67%, les bilingues en ont produit 47.62%, et les anglophones 41.98%.

Il se peut aussi que certains phonèmes soient plus fréquents dans certaines langues ce qui en faciliteraient l'acquisition et la prononciation, notamment en liaison. Monnin, Loevenbruck et Beckman (2007) ont mis en évidence le fait que lorsque des phonèmes sont moins fréquents dans une langue, les jeunes enfants les apprennent et les maîtrisent plus tard. Si les phonèmes les plus faciles à prononcer d'un point de vue articulatoire sont appris à peu près en même temps par tous les enfants du même âge, indépendamment de la L1, l'ordre d'acquisition des phonèmes un peu plus complexes à prononcer dépend de la fréquence de ces phonèmes dans la langue des apprenants.

On constate dans les données des tableaux 4.42., 4.43., 4.44 et 4.45., que la prononciation semble varier en fonction des consonnes de liaison, ce qui pourrait être dû à l'influence de celles qui sont présentes dans la langue maternelle des locuteurs. Les anglophones produisent les liaisons par ordre décroissant avec le /n/ (67.28%), le /z/ (63.23%), le /R/ (41.98%), le /v/ (40.74%), et enfin le /t/ (35.45%). Les bilingues produisent les liaisons par ordre décroissant avec le /z/ (73.47%), le /n/ (67.86%), le /R/ (47.62%), le /t/ (46.94 %) et enfin le /v/ (28.57%). Le locuteur chinois produit les liaisons par ordre décroissant avec le /n/ (83.33%), le /R/ (66.67%), le /z/ (64.29%), le /t/ 14.29 %) et enfin le /v/ (0%). Les locuteurs des autres langues produisent les liaisons par ordre décroissant avec le /R/ (83.33%), le /z/ (76.79%), le /n/ (62.50%), le /t/ (50 %) et enfin le /v/ (0%).

Nous voyons donc que la prononciation des consonnes de liaison semble dépendre, au moins en partie, de la langue maternelle, et ainsi faciliter ou limiter les productions de liaisons.

LO	T1			T2			T3		
	nb	tot	%	nb	tot	%	nb	tot	%
z	478	756	<b>63.23</b>	524	756	<b>69.31</b>	533	756	<b>70.50</b>
n	218	324	<b>67.28</b>	233	324	<b>71.91</b>	236	324	<b>72.84</b>
r	34	81	<b>41.98</b>	39	81	<b>48.15</b>	37	81	<b>45.68</b>
t	67	189	<b>35.45</b>	88	189	<b>46.56</b>	95	189	<b>50.26</b>
v	11	27	<b>40.74</b>	15	27	<b>55.56</b>	14	27	<b>51.85</b>
tot	808	1377	<b>58.68</b>	899	1377	<b>65.29</b>	915	1377	<b>66.45</b>

**Tableau 4.42. Productions de liaisons obligatoires – consonnes de liaison - anglophones**

LO	T1			T2			T3		
	nb	tot	%	nb	tot	%	nb	tot	%
z	144	196	<b>73.47</b>	166	196	<b>84.69</b>	165	196	<b>84.18</b>
n	57	84	<b>67.86</b>	69	84	<b>82.14</b>	64	84	<b>76.19</b>
r	10	21	<b>47.62</b>	14	21	<b>66.67</b>	12	21	<b>57.14</b>
t	23	49	<b>46.94</b>	33	49	<b>67.35</b>	31	49	<b>63.27</b>
v	2	7	<b>28.57</b>	2	7	<b>28.57</b>	2	7	<b>28.57</b>
tot	236	357	<b>66.11</b>	284	357	<b>79.55</b>	274	357	<b>76.75</b>

**Tableau 4.43. Productions de liaisons obligatoires – consonnes de liaison – bilingues**

LO	T1			T2			T3		
	nb	tot	%	nb	tot	%	nb	tot	%
z	43	56	<b>76.79</b>	50	56	<b>89.29</b>	46	56	<b>82.14</b>
n	15	24	<b>62.50</b>	20	24	<b>83.33</b>	21	24	<b>87.50</b>
r	5	6	<b>83.33</b>	6	6	<b>100.00</b>	3	6	<b>50.00</b>
t	7	14	<b>50.00</b>	11	14	<b>78.57</b>	12	14	<b>85.71</b>
v	0	2	<b>0.00</b>	1	2	<b>50.00</b>	2	2	<b>100.00</b>
tot	70	102	<b>68.63</b>	88	102	<b>86.27</b>	84	102	<b>82.35</b>

**Tableau 4.44. Productions de liaisons obligatoires – consonnes de liaison – allemand et arabe**

LO	T1			T2			T3		
	nb	tot	%	nb	tot	%	nb	tot	%
z	18	28	<b>64.29</b>	20	28	<b>71.43</b>	24	28	<b>85.71</b>
n	10	12	<b>83.33</b>	10	12	<b>83.33</b>	10	12	<b>83.33</b>
r	2	3	<b>66.67</b>	2	3	<b>66.67</b>	1	3	<b>33.33</b>
t	1	7	<b>14.29</b>	1	7	<b>14.29</b>	3	7	<b>42.86</b>
v	0	1	<b>0.00</b>	0	1	<b>0.00</b>	0	1	<b>0.00</b>
tot	31	51	<b>60.78</b>	33	51	<b>64.71</b>	38	51	<b>74.51</b>

**Tableau 4.45. Productions de liaisons obligatoires – consonnes de liaison – cantonnais**

Les productions de liaisons selon les consonnes de liaison semblent être l'un des éléments qui différencie le plus les natifs des non-natifs. Comme nous le voyons dans le

tableau 4.46., les francophones ne traitent pas les consonnes de liaison de la même façon. On voit déjà des différences entre les francophones majoritaires et minoritaires, mais les différences avec les locuteurs non francophones (tableau 4.47.) sont vraiment flagrantes.

tous francophones									
consonne	franc. Maj			franc. Min			total		
	N		%	N		%	N		%
<b>v</b>	12	12	<b>100</b>	8	8	<b>100</b>	20	20	<b>100</b>
<b>n</b>	144	144	<b>100</b>	89	96	<b>92,7</b>	233	240	<b>97,1</b>
<b>z</b>	331	336	<b>98,5</b>	214	224	<b>95,5</b>	545	560	<b>97,3</b>
<b>t</b>	70	84	<b>83,3</b>	32	56	<b>57,1</b>	102	140	<b>72,9</b>
<b>r</b>	28	36	<b>77,8</b>	6	24	<b>25</b>	34	60	<b>56,7</b>
<b>tot.</b>	585	612	<b>95,6</b>	349	408	<b>85,5</b>	934	1020	<b>91,6</b>

Tableau 4.46. Productions de liaisons obligatoires – consonnes de liaison – tous francophones

cons.	francophones						anglophones								
	Maj			Min			T 1			T 2			T 3		
LO	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%
<b>v</b>	12	12	<b>100</b>	8	8	<b>100</b>	13	37	<b>35,1</b>	18	37	<b>48,6</b>	18	37	<b>49</b>
<b>n</b>	144	144	<b>100</b>	89	96	<b>92,7</b>	683	1036	<b>65,9</b>	760	1036	<b>73,4</b>	768	1036	<b>74</b>
<b>z</b>	331	336	<b>98,5</b>	214	224	<b>95,5</b>	300	444	<b>67,6</b>	332	444	<b>74,8</b>	331	444	<b>75</b>
<b>t</b>	70	84	<b>83,3</b>	32	56	<b>57,1</b>	98	259	<b>37,8</b>	133	259	<b>51,4</b>	141	259	<b>54</b>
<b>r</b>	28	36	<b>77,8</b>	6	24	<b>25</b>	51	111	<b>45,9</b>	61	111	<b>55</b>	53	111	<b>48</b>
<b>tot.</b>	585	612	<b>95,6</b>	349	408	<b>85,5</b>	1145	1887	<b>60,7</b>	1304	1887	<b>69,1</b>	1311	1887	<b>70</b>

Tableau 4.47. Liaisons obligatoires et consonne de liaison, tous francophones et anglophones

La langue maternelle semble jouer un rôle sur le traitement des consonnes de liaison, cela se remarque déjà chez les francophones minoritaires qui subiraient l'influence d'une langue maternelle pour les Africains et d'une langue majoritaire pour les Franco-ontariens, en particulier avec le /t/ et le /R/. L'influence des langues parlées, qu'elles soient maternelles ou non influencerait le traitement des consonnes de liaison. Alors que les francophones minoritaires traitent pour la plupart (à l'exception du /n/) les consonnes de liaison selon la même hiérarchie que les francophones majoritaires, mais font beaucoup moins de liaisons avec le /t/ et le /R/, ce n'est pas le cas pour les autres groupes qui semblent traiter les consonnes de liaison, en partie selon l'influence de leur(s) langue(s) maternelle(s) et des langues parlées.



### 2.1.3. Résultats par catégories grammaticales

Si l'on compare les productions de liaisons par catégories grammaticales des francophones majoritaires et minoritaires et les productions des anglophones on observe d'énormes différences (tableaux 4.48., 4.49. et 4.50.). Les liaisons sont réalisées de façon systématique (à 100%) par tous les francophones pour certaines catégories grammaticales (les plus fréquentes) : dans les contextes [art. déf. + nom], [adj. dém. + nom] et [adj. num. + nom].

D'autres catégories sont réalisées de manière systématique ou quasi systématiques (entre 95 et 100%, selon les groupes de francophones) : [art. ind. + nom], [adj. ind. + nom] [prép. + art. + nom], [adj. pos. + nom] et [adj. num. + chiffre]. Enfin, les liaisons sont réalisées en plus faible proportion par les groupes majoritaires et minoritaires dans trois catégories : [adj. qual. + nom], [nom + adj.] et [adj. ord. + nom]. Les tendances sont globalement similaires pour les francophones majoritaires et minoritaires : plus les structures syntaxiques sont fréquentes (déterminant + nom) plus elles sont réalisées en plus grande proportion.

Les francophones majoritaires et minoritaires se basent donc sur les catégories syntaxiques pour traiter les liaisons. Selon Ollivier et Beaudoin (2010 : 3) : « les déterminants [...] peuvent indiquer la possession, la démonstration, la négation et certaines autres caractéristiques. Il existe plusieurs types de déterminants. Les plus importants sont les articles définis et indéfinis, les déterminants possessifs et les déterminants démonstratifs [...] ».

Les auteurs remarquent que parmi les déterminants, certains sont plus « importants », ce que nous interprétons comme voulant dire « plus fréquemment employés » (d'où l'importance de les connaître). Les déterminants que les auteurs estiment être les plus « importants » ou plus « fréquents », sont précisément ceux dont les liaisons sont produites de façon systématique ou quasi systématique par les locuteurs francophones (tableaux 4.48., 4.49.). Les francophones minoritaires observent la même tendance, c'est-à-dire qu'ils

observent une sorte de hiérarchie dans le traitement des liaisons, hiérarchie correspondant à la fréquence syntaxique de ces structures, bien que les liaisons soient réalisées en plus faible proportion.

Par contre, les anglophones n'observent pas ce traitement de la liaison par catégorie syntaxique, d'après la fréquence des structures. Tout d'abord, aucune catégorie syntaxique n'est réalisée à plus de 95%. Ceci suggère que les liaisons sont traitées plus au coup par coup que de façon systématique d'après des critères syntaxiques.

À l'intérieur d'une même catégorie syntaxique, on observe des différences. Par exemple, la liaison avec *nos amis* a été réalisée à 100% au temps 3, alors qu'avec *mon ami* ou *son école* elle a été réalisée à 94.6%. La liaison avec *leurs enfants* a été réalisée à 91.9%, avec *ton enfant* à 89.2%, avec *mes ouvriers* à 86.5% et avec *tes assistants* à 83.8%. Par contre, les liaisons ont été réalisées à 100% par les francophones majoritaires dans tous ces contextes (sauf une liaison non réalisée pour *mes ouvriers*).

On observe aussi que les contextes syntaxiques qui ont été les moins réalisés par les francophones majoritaires tels que [nom + adj.] (77.8%) ont été réalisés en plus grande proportion (66.7% au temps 3) par les anglophones que des contextes tels que [adj. ind. + nom] (58.1%) alors que ce contexte [adj. ind. + nom] a été réalisé à 100% par les francophones majoritaires. Cela indique que les anglophones n'ont pas les mêmes « priorités syntaxiques ».

Par contre, il est surprenant que les anglophones produisent un plus grand nombre de liaisons dans certaines catégories syntaxiques que les francophones minoritaires. Les anglophones produisent par exemple 33.8% de liaisons dans la catégorie [adj. ord. + nom] au temps 3 vs 25% pour les francophones minoritaires, et 66.7% dans la catégorie [nom + adj.] vs 54.2% pour les francophones minoritaires.

Ceci indique que les anglophones ne semblent pas avoir dans leur banque de données un système hiérarchique selon lequel certaines catégories sont plus « importantes » que d'autres (catégories basées sur la fréquence syntaxique). Ainsi, les anglophones peuvent ne pas faire les liaisons les plus « obligatoires » des contextes [art. déf. + nom] ou [art. ind. + nom], et faire des liaisons moins « obligatoires » telles que [adj. num. + nom].

Les expressions figées (tableau 4.50.) ont été réalisées par les anglophones au temps 3 dans une proportion similaire aux francophones minoritaires (72.1% vs 72.9%). Notons que la liaison avec *États-Unis* a été réalisée à 97.3% pour les anglophones ce qui est très élevé. Notons également que ce même mot est probablement plus utilisé au Canada, qu'il ne l'est en France ou en Belgique<sup>44</sup>, dans la mesure où les États-Unis sont le pays voisin du Canada et le plus grand partenaire économique.

Le même phénomène se retrouve avec *accent aigu* qui a été réalisé par les anglophones à 67.6%, soit presque deux fois plus qu'au temps 1, alors qu'il n'a été réalisé qu'à 58.3% par les francophones majoritaires. Là aussi, ce mot est peut-être plus employé pour les apprenants du français L2 que pour certains francophones (autres que les enseignants du français) qui ont rarement à épeler des mots et à utiliser ce mot. En effet, plusieurs francophones de notre étude n'enseignent pas le français et vivent dans un milieu exclusivement anglophone. Précisons aussi que plusieurs expressions figées étaient présentes dans la petite leçon, dont *États-Unis* et *accent aigu*. Il est possible que les étudiants se soient souvenus de ces expressions déjà connues et utilisées en classe à d'autres occasions.

Par contre, *Moyen-âge* a été réalisé à 86.5% par les anglophones et à 100% par les francophones majoritaires. *Fait accompli* a été réalisé à 35.1% par les anglophones et à 83.3% par les francophones majoritaires. Les francophones utilisent sans doute plus ces expressions

---

<sup>44</sup> Au moment de l'étude, certains locuteurs français ou africains ne restaient au Canada que pendant une année d'échange ; d'autres Français de l'étude habitaient au Canada depuis seulement quelques années ; enfin, l'apprentissage du français dans un autre pays aurait pour résultat de parler une variété dialectale différente du français canadien mais aussi d'avoir entendu certains mots ou expressions plus ou moins souvent.

dans leur conversation quotidienne que ne les entendraient et utiliseraient les apprenants en salle de classe.

À ce stade de leur développement langagier, les anglophones semblent apprendre les liaisons au coup par coup selon la fréquence lexicale des séquences et semblent ne pas traiter les liaisons de façon systématique d'un point de vue syntaxique.

liaisons obligatoires	franc						ang								
	francophones			min			T 1			T 2			T 3		
structure grammaticale	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%
(2) article déf.+ nom	24	24	100	16	16	100	50	74	67,6	52	74	70,3	55	74	74,3
(1) Adj. Dém.+ nom	12	12	100	8	8	100	26	37	70,3	31	37	83,8	30	37	81,1
(7) Adj.Num. + nom	84	84	100	56	56	100	178	259	68,7	199	259	76,8	195	259	75,3
(1) prép. + nom	12	12	100	8	8	100	27	37	73	31	37	83,8	28	37	75,7
(5) art. ind.+ nom	60	60	100	39	40	97,5	114	185	61,6	135	185	73	136	185	73,5
(2) Adj. Ind.+ nom	24	24	100	15	16	93,8	37	74	50	44	74	59,5	43	74	58,1
(3) prép. + art.+ nom	36	36	100	23	24	95,8	77	111	69,4	79	111	71,2	85	111	76,6
(7) Adj. pos. + nom	83	84	98,8	56	56	100	222	259	85,7	233	259	90	237	259	91,5
(2) Adj.Num.+ chif.	23	24	95,8	16	16	100	48	74	64,9	54	74	73	65	74	87,8
(16) Adj. Qual.+ nom	182	192	94,8	95	128	74,2	289	592	48,8	344	592	58,1	338	592	57,1
(3) Nom + adjectif	28	36	77,8	13	24	54,2	56	111	50,5	76	111	68,5	74	111	66,7
(2) Adj. Ord. + nom	17	24	70,8	4	16	25	21	74	28,4	26	74	35,1	25	74	33,8
<b>total</b>	585	612	95,6	349	408	85,5	1145	1887	60,7	1304	1887	69,1	1311	1887	69,5

Tableau 4.48. LO par catégories grammaticales, francophones majoritaires, minoritaires et anglophones

toutes occurrences	FMaj			T 1			T 2			T 3		
liaisons obligatoires	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%
7. nos amis	12	12	100	35	37	94,6	37	37	100	37	37	100
36. États-Unis	11	12	91,7	33	37	89,2	36	37	97	36	37	97,3
2. les affaires	12	12	100	34	37	91,9	34	37	92	35	37	94,6
3. des étudiants	12	12	100	31	37	83,8	33	37	89	35	37	94,6
11. plein air	12	12	100	34	37	91,9	34	37	92	35	37	94,6
12. mon avis	12	12	100	36	37	97,3	35	37	95	35	37	94,6
15. son école	12	12	100	35	37	94,6	35	37	95	35	37	94,6
45. trois arbres	12	12	100	33	37	89,2	35	37	95	35	37	94,6
6. leurs enfants	12	12	100	30	37	81,1	32	37	87	34	37	91,9
8. bon ami	12	12	100	33	37	89,2	35	37	95	34	37	91,9
40. des histoires	12	12	100	32	37	86,5	33	37	89	34	37	91,9
42. Deux amis	12	12	100	33	37	89,2	36	37	97	34	37	91,9
13. ton enfant	12	12	100	31	37	83,8	31	37	84	33	37	89,2
46. vingt trois	11	12	91,7	24	37	64,9	27	37	73	33	37	89,2
5. mes ouvriers	11	12	91,7	26	37	70,3	31	37	84	32	37	86,5
9. moyen-âge	12	12	100	33	37	89,2	34	37	92	32	37	86,5

48. vingt cinq	12	12	100	24	37	64.9	27	37	73	32	37	86.5
4. Tes assistants	12	12	100	29	37	78.4	32	37	87	31	37	83.8
28. aux yeux	12	12	100	27	37	73	28	37	76	31	37	83.8
47. trois ans	12	12	100	30	37	81.1	31	37	84	31	37	83.8
41. ces hôpitaux	12	12	100	26	37	70.3	31	37	84	30	37	81.1
1. aux invités	12	12	100	23	37	62.2	23	37	62	29	37	78.4
10. certain âge	12	12	100	27	37	73	31	37	84	29	37	78.4
50. cent ans	12	12	100	25	37	67.6	28	37	76	29	37	78.4
33. sous-entendu	12	12	100	27	37	73	31	37	84	28	37	75.7
32. faux-airs	11	12	91.7	23	37	62.2	28	37	76	26	37	70.3
30. beaux All.	12	12	100	20	37	54.1	21	37	57	25	37	67.6
35. accent aigu	7	12	58.3	14	37	37.8	25	37	68	25	37	67.6
51. dans un	12	12	100	27	37	73	28	37	76	25	37	67.6
19. dernier étage	12	12	100	22	37	59.5	26	37	70	24	37	64.9
21. nouveaux étud.	12	12	100	22	37	59.5	25	37	68	24	37	64.9
34. un accent	12	12	100	20	37	54.1	23	37	62	24	37	64.9
37. un hôpital	12	12	100	16	37	43.2	24	37	65	24	37	64.9
43. dix ans	12	12	100	24	37	64.9	28	37	76	24	37	64.9
44. six heures	12	12	100	20	37	54.1	23	37	62	24	37	64.9
16. premier amour	11	12	91.7	20	37	54.1	22	37	60	22	37	59.5
26. belles accl.	11	12	91.7	13	37	35.1	18	37	49	21	37	56.8
24. vieux écrits	12	12	100	18	37	48.6	21	37	57	20	37	54.1
25. nouvelles idées	11	12	91.7	17	37	45.9	18	37	49	20	37	54.1
29. les oies	12	12	100	16	37	43.2	18	37	49	20	37	54.1
27. un oiseau	12	12	100	15	37	40.5	22	37	60	19	37	51.4
39. vieilles éponges	12	12	100	10	37	27	17	37	46	18	37	48.6
49. neuf heures	12	12	100	13	37	35.1	18	37	49	18	37	48.6
20. ancien édifice	12	12	100	10	37	27	15	37	41	17	37	45.9
38. vieux habits	12	12	100	15	37	40.5	20	37	54	17	37	45.9
14. aucun effort	12	12	100	10	37	27	13	37	35	14	37	37.8
17. fait accompli	10	12	83.3	9	37	24.3	15	37	41	13	37	35.1
31. folles aventures	12	12	100	9	37	24.1	12	37	32	12	37	32.4
18. léger ennui	5	12	41.7	9	37	24.3	13	37	35	7	37	18.9
23. grand immeuble	12	12	100	1	37	2.7	7	37	19	6	37	16.2
22. second étage	6	12	50	1	37	2.7	4	37	11	3	37	8.1
Total	585	612	95.6	1145	1887	60.7	1304	1887	69	1311	1887	69.5

Tableau 4.49. Liaisons obligatoires par séquences. Francophones majoritaires et anglophones

LO	tous francophones						anglophones								
	fr	maj		fr	min		T1			T2			T3		
expres.	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%
figées	63	72	87.5	35	48	72.9	139	222	62.6	169	222	76.1	160	222	72.1

Tableau 4.50. Expressions figées (liaisons obligatoires). Tous francophones et anglophones

#### 2.1.4. Résultats par catégories fréquentielles

Nous avons établi en première partie de ce chapitre que les réalisations des liaisons et des enchaînements chez les francophones varient selon la fréquence du mot1, du mot2 et selon la fréquence de cooccurrence du mot1 et du mot2.

Les résultats de l'analyse de la fréquence du mot1 avec *Vocabprofil* (tableau 4.51.) n'indiquent pas clairement cette tendance pour les anglophones. Les anglophones font plus de liaisons (37.8% au temps 1) avec les mots de fréquence 1 qu'avec les mots de fréquence 2 (24.3%). Les résultats de l'analyse de la fréquence du mot1 avec *Vocabprofil* sont mitigés pour les francophones minoritaires.

Par contre, l'analyse de la fréquence du mot1 avec *Lexique* (tableau 4.52.) donne des résultats similaires pour les francophones majoritaires, les francophones minoritaires et les anglophones. Les francophones majoritaires font plus de liaisons avec les mots très fréquents (99.5%) qu'avec les mots moyennement fréquents (98%) et ils font plus de liaisons avec les mots moyennement fréquents (98%) qu'avec les mots peu fréquents (85.4%). De même, les francophones minoritaires font plus de liaisons avec les mots très fréquents (98.6%) qu'avec les mots moyennement fréquents (81%) et ils font plus de liaisons avec les mots moyennement fréquents (81%) qu'avec les mots peu fréquents (74%).

Quant aux anglophones, ils font eux aussi plus de liaisons avec les mots très fréquents (74.2% au temps 1) qu'avec les mots moyennement fréquents (59.8%) ils font plus de liaisons avec les mots moyennement fréquents (59.8%) qu'avec les mots peu fréquents (41.9%). On retrouve cette même tendance au temps 2 et au temps 3 (tableau 4.52.) Ainsi, la fréquence des mots1 est un facteur majeur dans le traitement de la liaison pour les trois groupes.

En outre, la proportion de liaisons réalisées semble être « décalée » pour chaque groupe. Les francophones minoritaires traitent la catégorie 1 (501+) approximativement comme les francophones majoritaires traitent les mots de la catégorie 2 (fréquence 41-500) et

traitent la catégorie 2 (fréquence 41-500) approximativement comme les francophones majoritaires traitent les mots de la catégorie 3 (fréquence 0-4).

De même, les anglophones traitent les mots de la catégorie 1, approximativement comme les francophones minoritaires traitent les mots de la catégorie 2, et les mots de la catégorie 2 approximativement comme les francophones minoritaires traitent les mots de la catégorie 3.

fr. M1	franc maj			franc min			tous apprenants								
	nb	tot	%	nb	tot	%	T 1			T 2			T 3		
Voc	nb	tot	%	nb	tot	%	Nb	tot	%	nb	tot	%	nb	tot	%
<b>3</b>	561	576	<b>97.4</b>	337	384	<b>87.8</b>	1113	1776	<b>62.7</b>	1254	1776	<b>70.6</b>	1267	1776	<b>71.3</b>
<b>2</b>	17	24	<b>70.8</b>	8	16	<b>50</b>	18	74	<b>24.3</b>	25	74	<b>33.8</b>	19	74	<b>25.7</b>
<b>1</b>	7	12	<b>58.3</b>	4	8	<b>50</b>	14	37	<b>37.8</b>	25	37	<b>67.6</b>	25	37	<b>67.6</b>
<b>0</b>	0	0	<b>0</b>	0	0	<b>0</b>	0	0	<b>0</b>	0	0	<b>0</b>	0	0	<b>0</b>
<b>Tot</b>	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>	1304	1887	<b>69.1</b>	1311	1887	<b>69.5</b>

**Tableau 4.51. Liaisons obligatoires et fréquence du mot1 avec *Vocabprofil*. Tous francophones et anglophones**

Fréquence	Mot1	franc maj			franc Min			tous apprenants								
		nb	tot	%	Nb	Tot	%	T 1			T 2			T 3		
Lexique	c	nb	tot	%	Nb	Tot	%	nb	tot	%	nb	tot	%	nb	tot	%
501+	ijklm	215	216	<b>99.5</b>	142	144	<b>98.6</b>	494	666	<b>74.2</b>	538	666	<b>80.8</b>	544	666	<b>81.7</b>
41-500	defgh	247	252	<b>98</b>	136	168	<b>81</b>	465	777	<b>59.8</b>	529	777	<b>68.1</b>	522	777	<b>67.2</b>
0-40	abc	123	144	<b>85.4</b>	71	96	<b>74</b>	186	444	<b>41.9</b>	237	444	<b>53.4</b>	245	444	<b>55.2</b>
total		585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>	1304	1887	<b>69.1</b>	1311	1887	<b>69.5</b>

**Tableau 4.52. Liaisons obligatoires et fréquence du mot1 avec *Lexique*. Tous francophones et anglophones**

En ce qui concerne les mots2 (tableaux 4.53. et 4.54.), le décalage entre les groupes est plus visible qu'avec les mots1. Alors qu'avec les mots1, le décalage est « d'une colonne », avec les mots2, le décalage est de « deux colonnes ». D'après les données fournies par *Vocabprofil*, les francophones minoritaires traitent les mots2 de fréquence 3 (les mots les plus fréquents), comme les francophones majoritaires traitent les mots de fréquence 1 (les mots moins fréquents). Les anglophones traitent en temps 1, les mots de fréquence 3, comme les francophones minoritaires traitent les mots2 de fréquence 0.

Les résultats de *Vocabprofil* indiquent que les francophones minoritaires comme les anglophones font plus de liaisons avec les mots2 de fréquence 0, qu'avec les mots2 de fréquence 1. Ceci n'est pas le cas pour les francophones majoritaires qui font plus de liaisons avec les mots de fréquence 1 qu'avec les mots de fréquence 0. Comme c'était le cas avec les mots1, avec *Lexique*, les anglophones traitent, en temps 1, les mots2 les plus fréquents (101+) (75%) comme les francophones minoritaires traitent les mots2 les moins fréquents (75.8%). En temps 3, les anglophones traitent les mots les plus fréquents (81.6%) comme les francophones minoritaires traitent les mots moyennement fréquents (79.1%).

Les trois groupes sont influencés par la fréquence des mots2, mais les groupes qui ont moins entendu les séquences produisent moins de liaisons dans toutes les catégories. C'est avec les mots1 et les mots2 les moins fréquents que les différences entre les groupes sont les plus grandes. Pour les francophones majoritaires, la différence entre les mots les plus fréquents (99.3% avec *Lexique*) et les moins fréquents (90.6% avec *Lexique*) n'est pas très élevée. Elle est deux fois plus élevée pour les francophones minoritaires (75.8% vs 95.7%). Elle est encore plus élevée pour les anglophones en temps 1 (47.3% pour les mots les moins fréquents et 75% pour les mots les plus fréquents, en temps 1). Cette différence de traitement entre les mots s'estompe légèrement au temps 3 (59% pour les mots les moins fréquents et 81.6% pour les mots les plus fréquents, en temps 3), bien que toujours très visible.

fr. M2	franc maj			franc min			tous apprenants								
	Nb	Tot	%	nb	tot	%	T 1			T 2			T 3		
Voc	Nb	Tot	%	nb	tot	%	Nb	tot	%	nb	tot	%	nb	tot	%
<b>3</b>	434	444	<b>97.7</b>	271	296	<b>91.6</b>	920	1369	<b>67.2</b>	1012	1369	<b>73.9</b>	1026	1369	<b>74.9</b>
<b>2</b>	70	72	<b>97.2</b>	40	48	<b>83.3</b>	129	222	<b>58.1</b>	151	222	<b>68</b>	147	222	<b>66.2</b>
<b>1</b>	22	24	<b>91.7</b>	5	16	<b>31.3</b>	19	74	<b>25.7</b>	30	74	<b>40.5</b>	30	74	<b>40.5</b>
<b>0</b>	59	72	<b>81.9</b>	33	48	<b>68.8</b>	77	222	<b>34.7</b>	111	222	<b>50</b>	108	222	<b>48.6</b>
<b>Tot</b>	585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>	1304	1887	<b>69.1</b>	1311	1887	<b>69.5</b>

Tableau 4.53. LO et fréquence du mot2 avec *Vocabprofil*. Tous francophones et anglophones



Fréquence	Mot2	franc maj			franc min			tous apprenants								
								T 1			T 2			T 3		
Lexique	C	nb	tot	%	nb	Tot	%	nb	tot	%	nb	tot	%	nb	tot	%
101+	ghim	274	276	<b>99.3</b>	176	184	<b>95.7</b>	638	851	<b>75</b>	689	851	<b>81</b>	694	851	<b>81.6</b>
21-80	cde	137	144	<b>95.1</b>	76	96	<b>79.2</b>	227	444	<b>51.1</b>	260	444	<b>58.6</b>	268	444	<b>60.4</b>
0-20	ab	174	192	<b>90.6</b>	97	128	<b>75.8</b>	280	592	<b>47.3</b>	355	592	<b>60</b>	349	592	<b>59</b>
Total		585	612	<b>95.6</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>	1304	1887	<b>69.1</b>	1311	1887	<b>69.5</b>

**Tableau 4.54. Liaisons obligatoires et fréquence du mot2 avec *Lexique*. Tous francophones et anglophones**

En ce qui concerne la fréquence de cooccurrence des mots1 et des mots2 (tableaux 4.55.), on remarque les mêmes phénomènes qu'avec la fréquence des mots isolés.

Les francophones majoritaires font plus de liaisons avec les séquences au taux de cooccurrence très élevé (99.24%) qu'avec les séquences au taux de cooccurrence moyennement élevé (98.89%), et font plus de liaisons avec les séquences au taux de cooccurrence moyennement élevé (98.89%) qu'avec les séquence au taux plus bas (92%).

De même, les francophones majoritaires font plus de liaisons avec les séquences au taux de cooccurrence très élevé (97.7%) qu'avec les séquences au taux de cooccurrence moyennement élevé (90%), et font plus de liaisons avec les séquences au taux de cooccurrence moyennement élevé (90%) qu'avec les séquence au taux plus bas (77.5%).

Enfin, les anglophones ont des tendances similaires. Ils font plus de liaisons avec les séquences au taux de cooccurrence très élevé (75.9%) qu'avec les séquences au taux de cooccurrence moyennement élevé (71.2%), et font plus de liaisons avec les séquences au taux de cooccurrence moyennement élevé (71.2%) qu'avec les séquence au taux plus bas (47.7%).

Tous les groupes traitent les liaisons en fonction de cette fréquence de cooccurrence, mais on observe aussi le même décalage. Les francophones minoritaires traitent les séquences au taux de cooccurrence le plus élevé (500+) (97.7%) comme les francophones majoritaires traitent les séquences au taux moyennement élevé (60-449). Les francophones minoritaires traitent les séquences au taux de cooccurrence moyennement élevé (60-449) (90%) comme les francophones majoritaires traitent les séquences au taux le plus faible (0-59) (92%). On

observe le même phénomène pour les anglophones, comparativement aux francophones minoritaires.

Fréquence cooccurr.		franc						tous apprenants								
		maj			min			T 1			T 2			T 3		
Lexique	c	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%	nb	tot	%
500+	KLM	131	132	<b>99.24</b>	86	88	<b>97.7</b>	309	407	<b>75.9</b>	336	407	<b>82.6</b>	331	407	<b>81.3</b>
60-449	EFGHI	178	180	<b>98.89</b>	108	120	<b>90</b>	395	555	<b>71.2</b>	430	555	<b>77.5</b>	446	555	<b>80.4</b>
0-59	ABCD	276	300	<b>92</b>	155	200	<b>77.5</b>	441	925	<b>47.7</b>	538	925	<b>58.2</b>	534	925	<b>57.7</b>
Total		585	612	<b>95.59</b>	349	408	<b>85.5</b>	1145	1887	<b>60.7</b>	1304	1887	<b>69.1</b>	1311	1887	<b>69.5</b>

**Tableau 4.55. Liaisons obligatoires et fréquence de cooccurrence des mot1-mot2 avec *Lexique*. Tous francophones et anglophones**

La fréquence du mot1, du mot2 et de cooccurrence du mot1 et du mot2, joue un rôle important pour tous les groupes qui semblent traiter les informations de la même façon, c'est-à-dire faire un nombre de liaisons proportionnel aux séquences entendues d'après leur fréquence. Nous supposons que puisque les francophones minoritaires et les anglophones ont entendu toutes les séquences moins souvent, ils traitent les séquences entendues avec un taux de réalisation de la liaison similaire au nombre de séquences entendues, donc plus bas.

Ceci explique le « décalage » et met en évidence le MBU. Les anglophones ont entendu les séquences les plus fréquentes peut-être le même nombre de fois que les francophones minoritaires ont entendu des séquences moyennement fréquentes, et approximativement le même nombre de fois que les francophones majoritaires ont entendu des séquences moins fréquentes. Si le nombre de séquences entendues et potentiellement pratiquées est similaire, les résultats de productions sont similaires pour tous les groupes. Ils produisent un pourcentage de liaisons similaires selon le nombre de séquences entendues.

Par contre, contrairement aux attentes, les apprenants semblent avoir fait plus de progrès avec les mots1 et mots2 moins fréquents, ce qu'on observe dans les tableaux 4.51-4.55. D'une part les séquences plus rares ne sont pas encore fossilisées ; d'autre part, il est plus facile de faire des progrès quand les chiffres de départ sont les plus bas.

### 2.1.5. Résultats par groupe expérimental

Si l'on sait que la fréquence d'écoute (Bybee, 2001, 2005) joue un rôle dans la production de liaisons, qu'en est-il des méthodes pédagogiques ? La répétition (la fréquence) est-elle plus importante que le fait de remarquer un principe (cf Schmidt, 2001), de comprendre un principe avec des explications théoriques (VanPatten, 2004) ou que de corriger les étudiants (Lyster, 1994, 1998, 2004. Lyster et Ranta, 1997) ?

Pour répondre à cette question, les sujets ont fait leur deuxième enregistrement après certaines activités, mais dans un ordre différent. Les quatre groupes ont pratiqué les trois activités : répétition, explication, correction. Le groupe 4 était le groupe témoin et a fait l'enregistrement en début et a fait les activités après l'enregistrement.

De plus, bien que tous les groupes aient eu toutes les activités, le temps passé pour chaque activité était légèrement différent. Ainsi, le groupe qui a reçu les répétitions en premier a eu un peu plus de répétitions au final que les autres groupes qui ont également eu des répétitions, mais en plus faible quantité.

En analysant les résultats par groupe (tableaux 4.57., 4.58., 4.59., 4.60.), nous pouvons observer plusieurs tendances. Au temps 2, tous les groupes ont fait des progrès, y compris le groupe 4 qui a fait des progrès avec les liaisons obligatoires et interdites bien qu'il n'ait pourtant pas encore reçu d'instruction formelle au moment de l'enregistrement. Les étudiants du groupe 4 avaient tout de même reçu des photocopies dans lesquelles il y avait des explications sur les liaisons et enchaînements, photocopies qu'ils étaient censés ne pas regarder avant l'enregistrement qu'ils ont peut être consultées.

Le groupe 1 qui a eu les trois types d'activités n'a pas fait plus de progrès que les autres groupes ; il en a même fait moins au temps 2. En ce qui concerne les liaisons obligatoires, le groupe 1 en a fait 6.7% de plus qu'au temps 1, le groupe 2 en a fait 11.8% de plus, le groupe 3 en a fait 7.4% de plus et le groupe 4 en a fait 7.9% de plus.

Autrement dit, le groupe qui a fait nettement plus de progrès avec les liaisons obligatoires que les autres groupes, c'est le groupe 2 qui n'a eu que des répétitions (la fréquence), ce qui reflète la fréquence des éléments qu'ils ont reçu dans l'input, suivi du groupe 4, puis du groupe 3. En outre, le groupe 2 a également fait de grands progrès avec les enchaînements, avec 48.5% au temps 1 et 60.8% au temps 2.

Nous ne pouvons pas vraiment expliquer pourquoi le groupe 4 a fait des progrès au temps 2 et pourquoi il a fait plus de progrès que le groupe 3 ou le groupe 1. Le simple fait de s'être familiarisé avec le texte a-t-il joué un rôle ? Les étudiants ont-ils prêté attention à la forme, regardé la feuille d'explication ? Le fait de répéter plusieurs fois le même texte, de se familiariser avec les mots, les structures et le texte a-t-il aidé les étudiants à produire plus de liaisons ?

Si l'on considère les progrès une semaine après le cours (temps 3), en ce qui concerne les liaisons obligatoires, on remarque que le groupe 1 en a fait 4.1% de plus, le groupe 2 en a fait 13.4% de plus, le groupe 3 en a fait 8.1% de plus et le groupe 4 en a fait 14.7% de plus. Il est toutefois difficile de tirer des conclusions certaines en ce qui concerne le groupe 4, car il n'y avait que deux étudiantes, non anglophones, qui étaient d'excellentes étudiantes pour commencer.

Si l'on considère seulement les trois premiers groupes, il s'avère que c'est le groupe dans lequel l'emphase a été sur les répétitions qui a fait le plus de progrès au temps 2 et au temps 3.

En ce qui concerne les enchaînements, entre le temps 1 et le temps 3, le groupe 1 en a fait 1.5 % de plus, le groupe 2 en a fait 14.6 % de plus, le groupe 3 en a fait 5.7 % de plus, le groupe 4 en a fait 15.4% de moins. C'est donc le groupe 2 qui a fait largement plus de progrès que les autres groupes pour les enchaînements.

Les liaisons facultatives et les liaisons interdites n'étaient pas beaucoup mentionnées dans la leçon, elles étaient mentionnées « au passage ». Il n'est donc pas surprenant que les pourcentages obtenus dans ces deux catégories ne soient guère plus élevés pour tous les groupes.

Toutefois, la production de liaisons interdites suggère un phénomène particulier. Afin de mieux interpréter les résultats des tableaux 4.57., 4.58., 4.59., 4.60. en ce qui concerne les liaisons interdites, nous avons mis dans le tableau 4.56. les pourcentages de liaisons **interdites réalisées** (non correctement produites). Nous remarquons dans le tableau 4.56. que les groupes 1 et 3 font moins d'erreurs après la leçon de phonétique avec les liaisons interdites. Le groupe 1 en produit 50% au temps 1, 37.9% au temps 2 et 41.4% au temps 3. Le groupe 3 en produit 48.6% au temps 1, 44.3% au temps 2 et 43.8% au temps 3. Le groupe 4 en produit 42.9% au temps 1, 39.3% au temps 2 et 42.9% au temps 3. Autrement dit, le groupe 1 et le groupe 3 font moins d'erreurs avec les liaisons interdites après la leçon. Le groupe 4 arrive en temps 3 au même niveau qu'au temps 1, restant donc au même niveau.

Par contre, le groupe 2 produit nettement plus de liaisons interdites après la leçon. Le groupe 2 en produit 34.3% au temps 1, 47.1% au temps 2 et 47.9% au temps 3. Rappelons que le groupe 2 était le groupe ayant eu plus de répétitions et au temps 2 et n'avait eu pratiquement aucune explication (2mn). Ce groupe 2 avait produit beaucoup moins de liaisons obligatoires au temps 1 que tous les autres groupes (tableau 4.58.) : seulement 52.5% tandis que les trois autres groupes en avait produit 63.7%. Les connaissances sur les liaisons du groupe 2 sont donc plus limitées que les autres groupes. Ainsi, la fréquence et la répétition pourraient avoir pour résultats que les locuteurs commencent à construire des représentations abstraites, surgénéralisent les règles et l'utilisation des liaisons et font des liaisons là où il ne faudrait pas en faire, n'en ayant pas encore maîtrisé les contraintes. Réaliser un plus grand nombre de liaisons interdites serait en fait le signe d'une acquisition en cours au même titre

qu'une baisse de production de liaisons interdites. On pourrait comparer cela aux enfants francophones qui font plus d'erreurs à un moment de leur acquisition (voir Chevrot, Dugua et Fayol, 2005), au moment où ils créent des représentations abstraites.

	temps 1	temps 2	temps 3
groupe 1	50	37.9	41.4
groupe 2	34.3	47.1	47.9
groupe 3	48.6	44.3	43.8
groupe 4	42.9	39.3	42.9

**Tableau 4.56. Production de liaisons interdites**

Le groupe 2 est celui qui a fait le plus de progrès avec les liaisons obligatoires (après le groupe 4 qui en a fait 14.7% de plus), en en produisant 13.4% de plus. Le groupe 1 en a produit 4.1% de plus, et le groupe 3, 8.1% de plus. Ce groupe 2 a également produit 13.6% de plus de liaisons interdites (fait 13.6% de plus d'erreurs), pourcentage qui se rapproche des progrès réalisés avec les liaisons obligatoires (13.4%).

C'est donc le groupe 2 qui a fait beaucoup plus de liaisons interdites qu'avant, mais qui a aussi fait beaucoup de liaisons obligatoires et enchaînements que les autres groupes. Ainsi, l'accroissement du nombre de liaisons interdites irait de pair avec l'accroissement de liaisons obligatoires et d'enchaînements, tout au moins à une étape d'acquisition, et serait un signe du développement langagier des apprenants.

Par contre, plus les apprenants sont à un stade avancé dans leur acquisition des liaisons, moins ils produisent de liaisons interdites. Ainsi, le groupe 4 qui au final a produit le plus de liaisons obligatoires (78.4%) et qui a réalisé 14.7% de plus de liaisons obligatoires au temps 3 qu'au temps 1 a produit le même pourcentage de liaisons interdites au temps 1 et au temps 3 (42.9%). De même, le groupe 2 qui a produit 8.1% de plus de liaisons obligatoires a produit 4.8% de moins de liaisons interdites et le groupe 1 qui a produit 4.1% de liaisons obligatoires, produit 8.6% de moins de liaisons interdites.

Le scénario développemental semble se rapprocher, d'une certaine façon, de celui des enfants francophones. En début d'acquisition les apprenants ne font que répéter des séquences entendues, puis ils commencent à construire des représentations abstraites, généralisant les règles aux liaisons interdites. Notons qu'au temps 1 de l'étude, le pourcentage élevé de liaisons interdites chez les quatre groupes, indique que les locuteurs ont commencé à développer des représentations abstraites.

Puis, la production de liaisons obligatoires augmente ainsi que la production de liaisons interdites à cause d'une surgénéralisation des règles. Enfin, la production de liaisons obligatoires augmente, tandis que la production de liaisons interdites baisse (groupe 4). Néanmoins, les pourcentages de liaisons obligatoires allant de 65.9% pour le groupe 3 à 78.4% pour le groupe 4, suggère que les liaisons ne sont pas encore acquises au niveau natif, ce qui explique en même temps une production encore très élevée de liaisons interdites allant de 41.4% pour le groupe 1 à 47.9% pour le groupe 2.

	<b>T 1</b>			<b>T 2</b>			<b>T 3</b>		
<b>groupe 1</b>	Nb	tot	%	nb	tot	%	nb	tot	%
enchaînements	69	130	<b>53.1</b>	72	130	<b>55.4</b>	71	130	<b>54.6</b>
L. obligatoires	325	510	<b>63.7</b>	359	510	<b>70.4</b>	346	510	<b>67.8</b>
L. facultatives	34	170	<b>20</b>	40	170	<b>23.5</b>	36	170	<b>21.2</b>
L. interdites	70	140	<b>50</b>	87	140	<b>62.1</b>	82	140	<b>58.6</b>
<b>total</b>	498	950	<b>52.4</b>	558	950	<b>58.7</b>	535	950	<b>56.3</b>

**Tableau 4.57. Évolution du traitement de la liaison après leçon. Groupe 1**

	<b>T 1</b>			<b>T 2</b>			<b>T 3</b>		
<b>groupe 2</b>	Nb	tot	%	nb	tot	%	nb	tot	%
enchaînements	63	130	<b>48.5</b>	79	130	<b>60.8</b>	82	130	<b>63.1</b>
L. obligatoires	268	510	<b>52.5</b>	328	510	<b>64.3</b>	336	510	<b>65.9</b>
L. facultatives	35	170	<b>20.6</b>	38	170	<b>22.4</b>	42	170	<b>24.7</b>
L. interdites	92	140	<b>65.7</b>	74	140	<b>52.9</b>	73	140	<b>52.1</b>
<b>total</b>	458	950	<b>48.2</b>	519	950	<b>54.6</b>	533	950	<b>56.1</b>

**Tableau 4.58. Évolution du traitement de la liaison après leçon. Groupe 2**

	T 1			T 2			T 3		
groupe 3	Nb	tot	%	nb	tot	%	nb	tot	%
enchaînements	111	195	<b>56.9</b>	133	195	<b>68.2</b>	122	195	<b>62.6</b>
L. obligatoires	487	765	<b>63.7</b>	544	765	<b>71.1</b>	549	765	<b>71.8</b>
L. facultatives	63	255	<b>24.7</b>	76	255	<b>29.8</b>	79	255	<b>31</b>
L. interdites	108	210	<b>51.4</b>	117	210	<b>55.7</b>	118	210	<b>56.2</b>
<b>total</b>	<b>769</b>	<b>1425</b>	<b>54</b>	<b>870</b>	<b>1425</b>	<b>61.1</b>	<b>868</b>	<b>1425</b>	<b>60.9</b>

**Tableau 4.59. Évolution du traitement de la liaison après leçon. Groupe 3**

	T 1			T 2			T 3		
groupe 4	Nb	tot	%	nb	tot	%	nb	tot	%
enchaînements	22	26	<b>84.6</b>	21	26	<b>80.8</b>	18	26	<b>69.2</b>
L. obligatoires	65	102	<b>63.7</b>	73	102	<b>71.6</b>	80	102	<b>78.4</b>
L. facultatives	9	34	<b>26.5</b>	7	34	<b>10.6</b>	10	34	<b>29.4</b>
L. interdites	16	28	<b>57.1</b>	17	28	<b>60.7</b>	16	28	<b>57.1</b>
<b>total</b>	<b>112</b>	<b>190</b>	<b>58.9</b>	<b>118</b>	<b>190</b>	<b>62.1</b>	<b>124</b>	<b>190</b>	<b>65.3</b>

**Tableau 4.60. Évolution du traitement de la liaison après leçon. Groupe 4**

Nous ne considérerons pas les progrès du groupe 4 dont le nombre de participants était trop peu nombreux pour en tirer des conclusions définitives. Il apparaît donc qu'au temps 2, c'est le groupe 2, celui dont l'emphase était sur la répétition, qui a fait le plus de progrès pour les liaisons obligatoires (+11.8%) et les enchaînements (+12.3%) et a aussi fait plus de liaisons interdites (+4.3%), par un processus d'analogie. C'est ensuite le groupe 3 qui au temps 2 a fait +7.4% de liaisons obligatoires et +11.3% d'enchaînements. Notons que le groupe 3 avait des répétitions et des corrections au moment de l'enregistrement.

Le groupe 1 qui au temps 1 avait pourtant réalisé autant de liaisons obligatoires que le groupe 3 (63.7%), produit au temps 2 légèrement moins de liaisons obligatoires (70.4%) que le groupe 3 (71.1%). Le groupe 1 avait pourtant bénéficié de 15 mn d'explications que le groupe 3 n'avait pas encore eues au moment de l'enregistrement. Ceci est compatible avec l'hypothèse selon laquelle la fréquence jouerait un rôle plus important que les corrections qui joueraient elles-mêmes un rôle plus important que les explications dans l'acquisition des liaisons.

Nous nous sommes efforcés de contrôler les effets potentiels d'autant de facteurs que possible, toutefois certains facteurs indépendants de notre volonté étaient tout de même



présents. Les résultats sont donc à considérer avec précaution, dans la mesure où plusieurs facteurs rentrent en jeu dans les progrès des étudiants, à commencer par leur niveau au départ. Ainsi, le groupe 2 était celui qui avait les résultats les plus faibles dans toutes les catégories avant la leçon. Il se peut que ce soit simplement le groupe le plus faible, ayant eu le moins de contact avec la langue avant la leçon qui ait bénéficié le plus de l'instruction.

Il se peut aussi que le type d'enseignement reçu à l'école ait joué un rôle dans la production et les progrès des étudiants. Par exemple, les étudiants étant le plus habitués à faire attention à la forme auraient pu faire plus attention à la forme que ceux habitués à se concentrer sur le message. C'est ce que nous allons étudier dans la prochaine section.

#### **2.1.6. Résultats par type de programme d'étude**

Le type d'enseignement reçu joue sans doute un rôle important dans la production de liaisons aux trois temps de notre étude. Il est utile de corrélérer les taux de réalisations de liaisons observés chez les locuteurs et les progrès réalisés au type d'enseignement reçu à l'école.

Tout d'abord, l'accès à la langue, l'input reçu, n'a pas été le même pour tous les groupes, et le type d'instruction non plus. Dans les écoles francophones, les élèves ont accès en bien plus grande proportion à la langue ; l'input est plus riche et plus varié étant donné que la majorité des enseignants sont francophones et que tous les cours se donnent en français dès la maternelle.

En immersion et en français intensif, seuls certains cours sont en français et l'input est réduit, comparativement aux écoles francophones, mais est plus grand qu'en français de base. L'input est réduit dans le programme de base, où il n'y a que 3-4 heures par semaine de cours de français et ce temps-là ne signifie pas que ceci est un temps d'exposition à la langue,

puisque'une partie des cours est en anglais. Ce type d'enseignement serait comparable dans le nombre d'heures comme dans le type d'enseignement aux cours d'anglais donnés en France.

Nous nous attendrions à ce que les groupes ayant eu une plus grande exposition à la langue produisent une plus grande proportion de liaisons, au moins de liaisons obligatoires. En effet, le sujet qui a fait le plus de liaisons obligatoires au début (80.4%) et à la fin (temps 3) (82.4%), est le sujet ayant étudié quelques années dans le système francophone, même si cela n'a pas été pendant toute sa scolarité (tableau 4.61.). Ces résultats sont très proches de ceux des francophones minoritaires (85.5%).

Soulignons au passage que les sujets africains auraient eu une exposition au français similaire à un élève en école francophone, le français étant la langue d'instruction parlée à l'école, mais peu utilisée en dehors de l'école. Un seul sujet a fréquenté une école francophone, parce que d'habitude les étudiants issus de ce système sont placés au moins en deuxième année, voire en troisième année (après le test de placement) et ne seraient donc pas dans ce cours de première année.

Le deuxième groupe qui a produit le plus de liaisons au temps 1 et au temps 3 (tableau 4.61.) est le groupe ayant étudié dans le système de français intensif, avec 68.6% au début et 77.1% à la fin. Rappelons que dans ce type d'enseignement, l'accent est sur la langue et qu'un seul cours est donné en français autre que le cours de français (cours de langue). C'est un système similaire au français de base, mais avec beaucoup plus d'heures de cours.

Vient ensuite, pour le temps 1, le groupe qui a étudié le français de base avec 62.4%. Il est assez surprenant que des personnes ayant entendu le français au maximum 3-4 heures par semaine (probablement moins, vu que de nombreuses explications sont données en anglais), aient produit plus de liaisons obligatoires que des personnes ayant entendu et pratiqué le français 25-30 heures par semaine dans des écoles d'immersion.

Plusieurs explications sont possibles. Tout d'abord, les personnes ayant étudié le français de base qui continuent le français à l'université sont celles qui ont terminé le français à l'école, c'est-à-dire, les meilleurs élèves (en français) à l'école. Les élèves qui étaient moins avancés et moins motivés ont généralement abandonné le français quand celui-ci n'était plus obligatoire.

Lorsque les élèves ne réussissent pas très bien en immersion, qu'ils changent d'école ou s'ils déménagent et n'ont pas accès à un programme d'immersion, ils iront probablement dans un autre système leur permettant de continuer le français (français de base ou français intensif). Ceci expliquerait pourquoi les sujets qui ont produit le moins de liaisons sont ceux qui ont fait leurs études dans plusieurs systèmes. Voyant qu'ils ne réussissaient pas très bien dans un système, ils sont peut-être allés tenter leur chance dans d'autres systèmes, espérant réussir mieux ailleurs.

Si l'on considère les progrès entre le temps 1 et le temps 3 (tableau 4.60.) par ordre décroissant de productions de liaisons obligatoires, les étudiants étant allés dans le système de français de base et d'immersion en ont fait + 15.5% ; ceux de français de base, immersion et français intensif + 13.7%, ceux d'immersion + 9.8%, ceux de français intensif + français de base 9.3%, ceux de français intensif + 8.5%, ceux de français de base + 7.7%, ceux de français intensif et immersion + 4.9% et ceux du système francophone et d'immersion + 2%. Notons cependant qu'il est difficile de faire des progrès quand on maîtrise déjà bien le système, comme c'est le cas pour l'étudiant étant allé dans le système francophone. De même, il est plus facile de faire des progrès lorsqu'on connaît très peu au début.

Tous les apprenants ont fait des progrès, quel que soit le système dont ils sont issus. Ceux qui ont fait le plus de progrès avec les liaisons obligatoires entre le temps 1 et le temps 3 sont ceux qui viennent du système d'immersion et du français intensif. Ceci pourrait s'expliquer par le fait qu'ils ont déjà entendu un bon nombre de séquences, ils ont déjà

bénéficié d'une fréquence élevée des formes testées dans l'input qu'ils ont reçu avec un plus grand nombre d'heures d'exposition à la langue et font donc le plus de progrès avec les séquences les plus fréquentes. Les apprenants ayant étudié le français de base ont aussi fait de grands progrès avec + 7.7%, ce qui est particulièrement remarquable étant donné leur exposition réduite à la langue.

Il est important de remarquer le phénomène de « régression » entre le temps 2 et le temps 3. Une seule semaine s'est écoulée entre le temps 2 et le temps 3. C'est-à-dire que la leçon est encore fraîche dans la mémoire des sujets. Pourtant, trois groupes ont produit légèrement moins de liaisons au temps 3 qu'au temps 2 : les sujets de français de base, d'immersion et de français intensif + français de base. Le sujet ayant été dans trois systèmes : français de base + immersion + français intensif n'a pas fait de progrès entre le temps 2 et le temps 3, bien qu'au moment du test en temps 2 il n'ait pas encore assisté à la totalité du cours.

Cela laisse présager que si l'instruction ne continue pas, en quelques mois ou en quelques semaines, les progrès apparents pourraient disparaître et les sujets pourraient revenir au point de départ, c'est-à-dire faire peu de liaisons. Les sujets venant de français de base et de français de base + français intensif, n'ont probablement pas assez entendu les séquences un nombre suffisant de fois pour avoir construit des représentations abstraites. Ainsi, ils ont répété les séquences comme ils les avaient entendues, sans pour autant véritablement internaliser les règles. S'ils avaient internalisé les règles, ils auraient produit un plus grand nombre de liaisons en temps 3 qu'en temps 2.

Par contre, les résultats des étudiants venant d'immersion soulèvent une autre question. Vu le nombre d'heures auxquelles ils ont été exposés au français, s'ils ont été exposés à des séquences bien formées (avec des liaisons et des enchaînements), ils devraient avoir entendu les séquences suffisamment de fois pour avoir internalisé les règles avant la leçon de prononciation et ils devraient donc avoir fait plus de liaisons que les sujets venant du

français de base. Or ce n'est pas le cas.

Lightbown et Spada (2006 : 193) expliquent un phénomène général selon lequel les élèves d'immersion ne font pas toujours des progrès dans les domaines qui n'entravent pas la compréhension et la communication :

[...] Teachers and researchers have also found, however, that the ability to understand the content and to function in classroom interaction does not insure that students will continue to improve in certain aspects of their second language. Especially in areas of accuracy on language features that do not usually interfere with meaning. Thus, for example, students can spend years in French immersion without achieving accuracy in marking nouns for gender or verbs for tense.

Le fait que les étudiants d'immersion fassent régulièrement des erreurs de genre suggère d'ailleurs que les mots n'ont pas été appris dans un groupe de mots, dans une séquence avec l'article correspondant, mais le mot a été appris sans l'article. Or, si les étudiants n'ont pas entendu suffisamment de fois les groupes de mots, ou n'ont pas fait assez attention pour avoir appris les articles correspondant aux noms (le genre des noms), il paraît peu probable qu'ils aient appris la prononciation des articles allant avec les noms, quand ils ne connaissent pas les articles eux-mêmes.

En outre, si les étudiants ne font pas attention aux « détails » de genre et d'accord, il est encore plus improbable qu'ils fassent attention aux « détails » de prononciation qui entravent encore moins la communication, dans la mesure où tous les apprenants de la classe comprennent l'accent anglophone et les erreurs phonétiques des locuteurs anglophones.

De plus, les étudiants venant d'immersion auraient dû aussi faire des progrès entre le temps 2 et le temps 3, ce qui n'a pas été le cas non plus. Cela suggère qu'ils ont peut-être fossilisé leur prononciation. Cela expliquerait pourquoi leur production de liaison est plus basse que celles des sujets de français de base, et pourquoi même s'ils ont fait des progrès, ils produisent légèrement moins de liaisons au temps 3 (70.1%) qu'au temps 2 (71%). Même s'ils ont momentanément répété les séquences telles qu'ils les ont entendues prononcées par l'enseignante, ils n'ont pas acquis le système de liaisons de façon systématique et définitive.

On ne peut pas tirer de conclusions précises sur le rôle des systèmes d’instruction vu que plusieurs facteurs rentrent en jeu, dont le temps passé en région francophone, l’exposition à des locuteurs natifs, la motivation des apprenants, etc. et surtout étant donné que le nombre de sujets dans chaque catégorie est restreint. On peut toutefois remarquer que plus les apprenants ont eu accès à la langue, meilleure est leur production et plus ils font des progrès, sauf ceux étant allés dans le système francophone qui maîtrisent déjà les liaisons obligatoires.

Temps 1				Temps 2				Temps 3			
école	nbLO	tot	%	école	nbLO	tot	%	école	nbLO	tot	%
imm et fr.	41	51	<b>80.4</b>	imm et fr.	39	51	<b>76.5</b>	imm et fr.	42	51	<b>82.4</b>
extendedF	105	153	<b>68.6</b>	extendedF	114	153	<b>74.5</b>	extendedF	118	153	<b>77.1</b>
core et ext	125	204	<b>61.3</b>	core + ext	149	204	<b>73</b>	core + ext	144	204	<b>70.6</b>
immersion	185	306	<b>60.5</b>	immersion	214	306	<b>69.9</b>	immersion	215	306	<b>70.3</b>
core	541	867	<b>62.4</b>	core	616	867	<b>71</b>	core	608	867	<b>70.1</b>
core + imm	79	153	<b>51.6</b>	core et imm	95	153	<b>62.1</b>	core et imm	103	153	<b>67.1</b>
c et im etex	24	51	<b>47.1</b>	c et im etex	31	51	<b>60.8</b>	c et im etex	31	51	<b>60.8</b>
ext + imm	45	102	<b>44.1</b>	ext + imm	46	102	<b>45.1</b>	ext + imm	50	102	<b>49</b>
total	1145	1887	<b>60.7</b>	total	1304	1887	<b>69.1</b>	total	1311	1887	<b>69.5</b>

**Tableau 4.61. Résultats par type d'école fréquentée - Liaisons obligatoires**

Si l’on considère les progrès accomplis dans toutes les catégories réunies, avec tous les types de sandhi (tableau 4.62.), on voit que ce sont surtout les étudiants issus du français de base, suivis de ceux du français intensif qui ont fait le plus de progrès. Au final, non seulement les sujets de ces deux systèmes ont fait le plus de progrès, mais ils produisent aussi plus de liaisons et enchaînements que ceux d’immersion. Ceux du français intensif en produisent 63.16%, ceux du français intensif et français de base 61.84%, ceux du français de base 59.57%, comparativement à ceux d’immersion (56.84%). Les différences sont particulièrement remarquables pour les enchaînements (tableau 4.63.). Ce sont les élèves d’immersion qui font le moins d’enchaînements au temps 3, ils en font moins qu’au temps 1.

Mis à part le sujet ayant passé quelque temps dans le système francophone qui a les meilleurs résultats dans toutes les catégories, ce sont ceux du français de base qui ont produit le plus grand nombre d’enchaînements, suivis de ceux du français intensif. C’est d’ailleurs dans cette catégorie que ces groupes ont fait le plus de progrès.

	temps 1	temps 3	Progrès
école	%	%	%
c + imm + ex	42.11	54.74	<b>12.63</b>
core + ext	51.58	61.84	<b>10.26</b>
core + imm	46.67	56.49	<b>9.82</b>
core	53.75	59.57	<b>5.82</b>
extendedF	57.89	63.16	<b>5.27</b>
immersion	51.93	56.84	<b>4.91</b>
ext + imm	40	43.16	<b>3.16</b>
imm + fr.	66.32	67.37	<b>1.05</b>
total	52.26	58.61	<b>6.35</b>

Tableau 4.62. Progrès tous types de sandhi

temps 1				temps 2				temps 3			
école	NbE	tot	%	école	NbE	tot	%	école	NbE	tot	%
imm + fr.	10	13	<b>76.9</b>	imm + fr.	7	13	<b>53.8</b>	imm + fr.	10	13	<b>76.9</b>
core + ext	28	52	<b>53.8</b>	core + ext	35	52	<b>67.3</b>	core + ext	34	52	<b>65.5</b>
core	132	221	<b>59.7</b>	core	154	221	<b>69.7</b>	core	143	221	<b>64.7</b>
c+imm+ex	4	13	<b>30.8</b>	c+imm+ex	9	13	<b>69.2</b>	c+imm+ex	8	13	<b>61.5</b>
extendedF	20	39	<b>51.3</b>	extendedF	22	39	<b>56.4</b>	extendedF	23	39	<b>59</b>
ext + imm	10	26	<b>38.5</b>	ext + imm	13	26	<b>50</b>	ext + imm	15	26	<b>57.7</b>
core + imm	17	39	<b>43.6</b>	core + imm	20	39	<b>51.3</b>	core + imm	21	39	<b>53.8</b>
immersion	44	78	<b>56.4</b>	immersion	45	78	<b>57.7</b>	immersion	39	78	<b>50</b>
total	265	481	<b>55.1</b>	total	305	481	<b>63.4</b>	total	293	481	<b>60.9</b>

Tableau 4.63. Résultats par type d'école fréquentée – enchaînements

Temps 1				Temps 2				Temps 3			
Ecole	NbLF	tot	%	Ecole	NbLF	tot	%	école	NbLF	tot	%
core+imm+ex	6	17	<b>35.3</b>	c+imm+ex	7	17	<b>41.2</b>	c+imm+ex	8	17	<b>47.1</b>
core + ext	16	68	<b>23.5</b>	core + ext	20	68	<b>29.4</b>	core + ext	24	68	<b>35.3</b>
extendedF	15	51	<b>29.4</b>	extendedF	12	51	<b>23.5</b>	extendedF	16	51	<b>31.4</b>
Core	63	289	<b>21.8</b>	Core	79	289	<b>27.3</b>	core	77	289	<b>26.6</b>
Immersion	24	102	<b>23.5</b>	immersion	26	102	<b>25.5</b>	immersion	25	102	<b>24.5</b>
imm + fr.	4	17	<b>23.5</b>	imm + fr.	3	17	<b>17.6</b>	imm + fr.	4	17	<b>23.5</b>
core + imm	11	51	<b>21.6</b>	core + imm	12	51	<b>23.5</b>	core + imm	11	51	<b>21.6</b>
ext + imm	2	34	<b>5.9</b>	ext + imm	2	34	<b>2.9</b>	ext + imm	2	34	<b>5.9</b>
Total LF	100	425	<b>22.4</b>	Total LF	118	425	<b>25.6</b>	Total LF	125	425	<b>27</b>

Tableau 4.64. Résultats par type d'école fréquentée - liaisons facultatives

Temps 1				Temps 2				Temps 3			
École	NbLI	tot	%	École	NbLI	tot	%	école	NbLI	tot	%
core + imm	26	42	<b>61.9</b>	core + imm	24	42	<b>57.1</b>	core + imm	26	42	<b>61.9</b>
core + ext	27	56	<b>48.2</b>	core + ext	30	56	<b>53.6</b>	core + ext	33	56	<b>58.9</b>
imm + fr.	8	14	<b>57.1</b>	imm + fr.	8	14	<b>57.1</b>	imm + fr.	8	14	<b>57.1</b>
Core	132	238	<b>55.5</b>	core	141	238	<b>59.2</b>	core	134	238	<b>56.3</b>
extendedF	25	42	<b>59.5</b>	extendedF	24	42	<b>57.1</b>	extendedF	23	42	<b>54.8</b>
Immersion	43	84	<b>51.2</b>	immersion	50	84	<b>59.5</b>	immersion	45	84	<b>53.6</b>
ext + imm	19	28	<b>67.9</b>	ext + imm	12	28	<b>42.9</b>	ext + imm	15	28	<b>53.6</b>
c + imm + ex	6	14	<b>42.9</b>	c + imm + ex	6	14	<b>42.9</b>	c + imm + ex	5	14	<b>35.7</b>
Total LI	286	518	<b>55.2</b>	total	295	518	<b>56.9</b>	total	289	518	<b>55.8</b>

Tableau 4.65. Résultats par type d'école fréquentée - liaisons interdites

temps 1				temps 2				temps 3			
école	LR	N	%	école	LR	N	%	école	LR	N	%
imm + fr.	63	95	<b>66.32</b>	imm + fr.	57	95	<b>60</b>	imm + fr.	64	95	<b>67.37</b>
extendedF	165	285	<b>57.89</b>	extendedF	172	285	<b>60.4</b>	extendedF	180	285	<b>63.16</b>
core + ext	196	380	<b>51.58</b>	core + ext	234	380	<b>61.6</b>	core + ext	235	380	<b>61.84</b>
core	868	1615	<b>53.75</b>	core	990	1615	<b>61.3</b>	core	962	1615	<b>59.57</b>
immersion	296	570	<b>51.93</b>	immersion	335	570	<b>58.8</b>	immersion	324	570	<b>56.84</b>
core + imm	133	285	<b>46.67</b>	core + imm	151	285	<b>53</b>	Core + imm	161	285	<b>56.49</b>
c+ imm +ex	40	95	<b>42.11</b>	c+ imm +ex	53	95	<b>55.8</b>	c+ imm +ex	52	95	<b>54.74</b>
ext + imm	76	190	<b>40</b>	ext + imm	73	190	<b>38.4</b>	ext + imm	82	190	<b>43.16</b>
total	1837	3515	<b>52.26</b>	total	2065	3515	<b>58.7</b>	total	2060	3515	<b>58.61</b>

**Tableau 4.66. Résultats par type d'école fréquentée - Tous types de sandhi**

### 2.1.7. Résultats par nombre d'années passées à étudier le français

On s'attendrait à ce que les sujets qui ont passé plus de temps à entendre, étudier et pratiquer le français, maîtrisent mieux le français, y compris les liaisons obligatoires. Les sujets ayant étudié le français 11-12 ans ou 13-14 ans devraient produire plus (sinon beaucoup plus) de liaisons que ceux qui l'ont étudié la moitié du temps (5-6 ans ou 7-8 ans). Ceci devrait d'autant plus être le cas que s'ils ont étudié le français aussi longtemps, cela signifie qu'ils ont commencé à étudier le français à un très jeune âge (en maternelle).

Or, si l'on regarde le tableau 4.67., classé par ordre décroissant de productions, on remarque que ce n'est pas le cas. La production de liaisons ne semble pas varier en fonction du nombre d'années étudiées. La tendance est presque inversée. Ceux qui ont étudié le français entre 7 et 8 ans ont de meilleurs résultats que ceux qui l'ont étudié 9-10, 11-12 et 13-14 ans et ceux qui l'ont étudié entre 5 et 6 ans ont de meilleurs résultats que tous les autres groupes, sauf ceux qui l'ont étudié 7-8 ans. Par contre, ceux qui ont fait le plus de progrès sont ceux qui l'ont étudié entre 9 et 10 ans et juste après, le groupe ayant étudié le français 11-12 ans.

Si la fréquence joue un rôle déterminant dans l'acquisition de la langue et ici l'acquisition des liaisons, cette fréquence ne joue pas de rôle dans la durée. Ainsi, c'est en début d'acquisition que cette fréquence semble jouer un rôle crucial. Rapidement, les sujets



acquièrent des habitudes et la fréquence dans le temps (à long terme) ne change pas leurs habitudes déjà établies, ce qui semble être le cas par exemple pour les élèves d’immersion.

D’autre part, les sujets qui ont étudié le français moins d’années sont souvent ceux qui ont étudié le français intensif ou le français de base. Il est possible que ces systèmes mettent d’avantage l’accent sur la forme et la précision de la langue dès le début contrairement aux écoles d’immersion où le français est utilisé pour enseigner d’autres cours et où l’emphase est davantage sur la fluidité que sur la précision.

temps 1			
Nb années	LR	N	%
7 et 8	237	380	62.4
5 et 6	222	380	58.4
9 et 10	679	1330	51.1
13 et 14	153	285	53.7
11 et 12	546	1140	47.9
total	1837	3515	<b>52.3</b>

temps 2			
Nb années	LR	N	%
7 et 8	257	380	67.6
5 et 6	276	380	72.6
9 et 10	773	1330	58.1
13 et 14	160	285	56.1
11 et 12	599	1140	52.5
total	2065	3515	<b>58.7</b>

temps 3			
Nb années	LR	N	%
7 et 8	245	380	64.5
5 et 6	238	380	62.6
9 et 10	791	1330	59.5
13 et 14	169	285	59.3
11 et 12	617	1140	54.1
total	2060	3515	<b>58.6</b>

**Tableau 4.67. Nombre d’années étudiées -Tous types de sandhi**

Par contre, les personnes ayant étudié le français moins d’années que les autres (7-8 ans et 5-6 ans) sont celles qui régressent au temps 3 par rapport au temps 2, ce qui suggère que l’apprentissage est incomplet ou que les sujets sont en cours d’apprentissage et qu’à plus long terme les sujets pourraient revenir à leurs habitudes de départ s’ils ne reçoivent pas la fréquence qui leur est nécessaire pour enraciner les structures. Il n’en demeure pas moins que les sujets ayant étudié le français entre 9 et 14 ans ont au temps 3 une production de liaisons toujours inférieure à celles des étudiants ayant étudié le français entre 5 et 6 ans.

### **2.1.8. Résultats par séjour en milieu francophone**

Là aussi, on s’attendrait à ce que les sujets qui ont passé du temps en milieu francophone aient plus entendu et pratiqué le français et produisent donc un plus grand nombre de liaisons obligatoires que les sujets qui ne sont jamais allés en milieu francophone.

Toutefois, ce n'est pas vraiment le cas (tableau 4.68. et 4.69.). Le sujet qui a passé plus d'une année en milieu francophone est le sujet qui a produit le moins de liaisons. Ceux qui ont passé entre 2-3 mois en milieu francophone en font moins que ceux qui ont passé moins de temps (1-2 semaines. 3-4 semaines ou 5-7 semaines). Les sujets qui n'y sont jamais allés, produisent au temps 1 plus de liaisons (61%) que ceux qui y ont passé 1-2 semaines, 2-3 mois ou même plus d'une année. Au temps 3, le temps passé en milieu francophone ne semble pas directement corrélé aux liaisons réalisées, même si, dans l'ensemble, les sujets qui ne sont jamais allés en milieu francophone produisent moins de liaisons que ceux qui y sont allés, à l'exception du sujet y ayant passé plus d'une année. En outre, les sujets qui ont fait le plus de progrès entre le temps 1 et le temps 3 sont ceux qui n'ont jamais passé de temps en milieu francophone.

Le sujet qui a passé entre 7 mois et un an a régressé entre le temps 1 et le temps 3. Cela peut s'expliquer par le fait que ce sujet qui avait les productions les plus élevées (78.4%) au temps 1, produisait les liaisons de façon intuitive, et la leçon l'a peut-être embrouillé. Réfléchissant aux règles et exceptions, le sujet a peut-être moins fait appel à son intuition et a tenté d'utiliser les règles vues en classe, ce qui l'a poussé au final à faire des erreurs et à en produire moins.

Nous pouvons donc en conclure que si les séjours en milieu francophone aident à développer le système de liaisons, d'une part il faut y avoir passé au moins trois semaines (ceux qui ont passé 1-2 semaines en font moins que ceux qui n'y sont jamais allés), d'autre part, passer du temps en milieu francophone ne permet pas de prédire des progrès en prononciation. Si tel était le cas, les sujets ayant passé plus de temps auraient de meilleurs résultats que ceux ayant passé moins de temps, qui eux-mêmes auraient de meilleurs résultats que ceux qui n'ont jamais passé de temps en milieu francophone.

Tout d’abord, passer du temps en milieu francophone, ne veut pas toujours dire que les sujets ont eu un input riche et varié et qu’ils ont pratiqué la langue. En effet, si le sujet ayant vécu à Hull était dans un milieu anglophone, l’exposition au français peut avoir été très limitée. Si des sujets partent en vacances avec leur famille, ils parleront tout le temps leur langue maternelle et le temps passé à entendre et pratiquer le français sera aussi très réduit.

En outre, si des personnes passent du temps en milieu francophone, mais restent chez des connaissances ou de la famille qui parle leur langue, alors ils n’auront pas besoin de parler le français. Enfin, si des personnes passent du temps en région francophone, au Québec, en France ou ailleurs, vu le grand nombre de francophones qui parlent l’anglais, il est possible que ces personnes puissent passer du temps en région francophone tout en communiquant uniquement en anglais. Passer du temps en région francophone n’est donc pas synonyme de parler, pratiquer et améliorer son français oral, notamment en ce qui concerne les liaisons.

Temps 1

Nb années	N	%
Jamais allé :	591	61
1 - 2 semaines	270	58.8
3 - 4 semaines :	104	68
5 - 7 semaines :	64	62.7
2 - 3 mois :	57	55.9
7 mois – 1 an :	40	78.4
1 an + :	19	37.3
total	1145	<b>60.7</b>

Temps 3

Nb années	N	%
Jamais allé :	650	67.1
1 - 2 semaines	327	71.2
3 - 4 semaines :	118	77.1
5 - 7 semaines :	79	77.5
2 - 3 mois :	72	70.6
7 mois – 1 an :	39	76.5
1 an + :	26	51
total	1311	<b>69.5</b>

**Tableau 4.68. Liaisons obligatoires et temps passé en pays/région francophone**

Temps 1

Temps passé	LOR	N	%
5 - 7 semaines	64	102	<b>62.7</b>
3 - 4 semaines	104	153	<b>68</b>
7 mois – 1 an	40	51	<b>78.4</b>
1 - 2 semaines	270	459	<b>58.8</b>
2 - 3 mois	57	102	<b>55.9</b>
Jamais allé	591	969	<b>61</b>
1 an +	19	51	<b>37.3</b>
total	1145	1887	<b>60.7</b>

Temps 3

Temps passé	LOR	N	%
5 - 7 semaines	79	102	<b>77.5</b>
3 - 4 semaines	118	153	<b>77.1</b>
7 mois – 1 an	39	51	<b>76.5</b>
1 - 2 semaines	327	459	<b>71.2</b>
2 - 3 mois	72	102	<b>70.6</b>
Jamais allé	650	969	<b>67.1</b>
1 an +	26	51	<b>51</b>
total	1311	1887	<b>69.5</b>

**Tableau 4.69. Liaisons obligatoires et temps passé en pays/région francophone**

### **2.1.9. Résultats par activités en français**

Nous avons souhaité évaluer cette caractéristique des apprenants parce que d'après le MBU, plus les sujets ont accès à la langue, plus ils ont de chances de maîtriser divers aspects de la langue et donc ici les liaisons.

La production des apprenants en temps 1 semble être directement liée au nombre d'activités (tableau 4.70.). Plus les sujets participent à des activités en français, plus ils produisent de liaisons. En temps 3, même s'ils ont fait des progrès, les sujets ne faisant pas d'activités ou faisant le moins d'activités (entre 1 et 2 activités) ont les scores les plus bas et les sujets faisant le plus d'activités ont les scores les plus hauts. Ceux qui font entre 3 et 8 activités semblent arriver à un niveau similaire au temps 3. C'est-à-dire qu'ils fassent 3, 4, 5, 6, 7, ou 8 activités ne semble pas faire beaucoup de différence dans leurs productions. Ceux qui font 7-8 activités ont les mêmes résultats que ceux qui font 3-4 activités.

Ainsi, les activités en français semblent être un facteur déterminant dans la production de liaisons. Ceci est assez logique puisque si les apprenants écoutent des chansons (et les chantent), écoutent la radio, regardent la télévision et parlent avec des francophones, ils sont beaucoup plus exposés à la langue, et leur input plus riche et plus varié leur a permis et leur permet de développer leur banque de données et de construire des représentations abstraites.

Ce facteur est d'ailleurs lié à d'autres facteurs tels que la motivation. La motivation est une variable qui a beaucoup été étudiée dans les recherches sur l'acquisition d'une L2 notamment dans les recherches de Dörnyei (1994), Dörnyei (2001), Gardner (1982), Gardner (2001). Gardner (2001) a montré que les apprenants motivés aiment apprendre une langue seconde, se fixent des buts et font plus d'efforts.

Si les apprenants sont motivés et veulent s'améliorer, ils se placeront dans les conditions qui leur permettent telles que ces activités en question. Une personne qui participe régulièrement à plus de neuf activités en français est très motivée pour apprendre le

français, et une personne qui ne participe à aucune activité ou seulement une ou deux (généralement les activités requises pour un cours de français) est une personne peu motivée.

Rappelons qu'au Canada, et plus précisément (en ce qui concerne notre étude) en Ontario, il est facile de trouver des activités gratuites en français. Il existe plusieurs clubs de français, un centre francophone pour tous les âges, des activités pour les parents et leurs jeunes enfants, un cinéclub en français, la possibilité d'emprunter des CD et des vidéos à la bibliothèque municipale ; il y a chaque année à l'université Western Ontario des étudiants français qui font leurs études dans des programmes d'échange, des assistants de laboratoire qui viennent de France, de nombreux Franco-ontariens, Québécois et autres francophones. La chaîne de télévision TV5 est gratuite (accessible sans payer le câble) et il existe plusieurs stations de radio en français.

Il est possible de faire du bénévolat dans les écoles, les centres francophones, les départements de français ou autres endroits où il y a des francophones. Tout le monde (en tout cas, tous les sujets de notre étude) a accès à l'internet où l'on peut trouver des chansons, les nouvelles, des films ou passages de films, des blogs et autres sites de discussions. Bref, il est facile et gratuit de trouver des activités en français pour tous ceux qui le souhaitent.

temps 1				temps 2				temps 3			
Nb activités	LOR	N	%	Nb activités	LOR	N	%	Nb activités	LOR	N	%
0 activ.	251	570	<b>44</b>	0 activ.	278	570	<b>48.8</b>	0 activ.	285	570	<b>50</b>
1-2 activ.	177	380	<b>46.6</b>	1-2 activ.	198	380	<b>52.1</b>	1-2 activ.	198	380	<b>52</b>
3-4 activ.	667	1235	<b>54</b>	3-4 activ.	732	1235	<b>59.3</b>	3-4 activ.	760	1235	<b>62</b>
5-6 activ.	469	855	<b>54.9</b>	5-6 activ.	535	855	<b>62.6</b>	5-6 activ.	512	855	<b>60</b>
7-8 activ.	214	380	<b>56.3</b>	7-8 activ.	254	380	<b>66.8</b>	7-8 activ.	236	380	<b>62</b>
9 + activ.	59	95	<b>62.1</b>	9 + activ.	68	95	<b>71.6</b>	9 + activ.	69	95	<b>73</b>
total	1837	3515	<b>52.3</b>	total	2065	3515	<b>58.7</b>	Total	2060	3515	<b>59</b>

**Tableau 4.70. Tous types de sandhi et nombre d'activités en français**

La motivation est liée au nombre d'activités en français ce qui permet d'augmenter l'input et l'interaction en français, conditions nécessaires à l'acquisition des liaisons.

### 2.1.10. Résultats selon auto-évaluation

Les sujets ont rempli un questionnaire dans lequel l'une des questions leur demandait de s'auto-évaluer en se donnant une note sur 100 pour sa prononciation. Il est utile de constater ici que d'après Hieke (1984), les liaisons et enchaînements sont assez représentatifs des connaissances phonétiques des sujets.

D'après nos résultats (tableau 4.71.), les étudiants semblent avoir une assez bonne idée de leur niveau de prononciation même s'ils peuvent légèrement se surestimer ou se sous-estimer. En effet, ceux qui ont produit le plus de liaisons se sont donné une note entre 85-90, ce qui est une excellente note. Ceux qui ont produit le moins de liaisons se sont donné moins de 60% ou entre 65 et 70, ce qui est une note faible. Les personnes qui se sont donné plus de 85% ont produit au temps 3 plus de 80% de liaisons obligatoires, ce qui est un excellent score.

Temps 1				Temps 3			
Note estimée	LOR	N	%	Note estimée	LOR	N	%
85-90	116	153	<b>75.8</b>	85-90	124	153	<b>81</b>
90-95	36	51	<b>70.6</b>	90-95	41	51	<b>80.4</b>
75-80	441	714	<b>61.8</b>	75-80	520	714	<b>72.8</b>
80-85	134	204	<b>65.7</b>	80-85	143	204	<b>70.1</b>
70-75	246	408	<b>60.3</b>	70-75	282	408	<b>69.1</b>
60-65	64	102	<b>62.7</b>	60-65	61	102	<b>59.8</b>
60-	25	51	<b>49</b>	60-	28	51	<b>54.9</b>
65-70	83	204	<b>40.7</b>	65-70	112	204	<b>54.9</b>
total	1145	1887	<b>60.7</b>	total	1311	1887	<b>69.5</b>

**Tableau 4.71. Niveau estimé en prononciation (auto-évaluation) et productions de liaisons obligatoires**

Les étudiants ont globalement une bonne idée de leur prononciation. Lorsqu'ils évaluent leur prononciation, ils pensent rarement à la production de liaisons, mais pensent plutôt à des éléments tels que la prononciation d'éléments segmentaux, la rapidité et la facilité d'élocution. Ainsi, si les apprenants se sont jugés faibles en prononciation et produisent peu de liaisons, il est probable que les liaisons soient entravées par des facteurs phonétiques autres que la liaison tels que le rythme, la production d'éléments segmentaux, la prononciation de consonnes finales, etc. Si les sujets ne savent pas quand et comment prononcer les consonnes

en fin de mots, ils vont forcément faire des erreurs avec les liaisons, par exemple prononcer les consonnes finales, mais ne pas les enchaîner.

Les liaisons sont intimement liées à d'autres éléments segmentaux et suprasegmentaux. Ainsi la liaison est un indice d'autres lacunes en prononciation et en cela reflète la prononciation des apprenants. Si un sujet fait 40% de liaisons obligatoires, il est peu probable que ce même sujet maîtrise tous les autres aspects phonétiques. Par contre, des sujets peuvent produire un grand nombre de liaisons tout en ne maîtrisant pas forcément tous les autres aspects phonétiques, mais au moins en maîtrisent-ils certains.

## Discussion

Les étudiants anglophones ne maîtrisent pas les liaisons au même niveau que les francophones. D'une part, ils produisent (au temps 1) 60.7% de liaisons obligatoires, résultats très éloignés des francophones majoritaires (95.6%) et des francophones minoritaires (85.5%) et ces résultats sont également éloignés des études de Thomas (2002-2004), Howard (2004) et Mastromonaco (2000). Ils produisent aussi moins d'enchaînements (55.1%) que les francophones majoritaires (78.2%), mais légèrement plus que les francophones minoritaires (54.8%)

Leurs productions indiquent qu'à ce stade de leur apprentissage (temps 1, temps 2, temps 3) ils ont appris et apprennent les liaisons au coup par coup, un peu comme les enfants francophones en stade initial d'apprentissage linguistique apprennent des éléments isolés sans être capables de généraliser ces règles de façon systématique à des éléments nouveaux, ou moins fréquents (Ellis, 2002 ; Tomasello, 2000a, 2000b ; Dugua, 2005).

En effet, ils font la liaison par exemple avec *nos amis* à 100%, mais ne font la liaison avec *mes ouvriers* qu'à 86.5%, alors qu'il s'agit pourtant du même contexte phonétique et syntaxique, ce qui suggère qu'ils ne généralisent pas encore les règles à des séquences non

entendues, ou moins entendues et qu'ils n'ont pas encore créé de catégories abstraites complètes leur permettant de maîtriser des règles. La raison en est qu'ils ont souvent entendu la séquence *nos amis (les amis. mes amis. tes amis...)* à l'école, tandis qu'ils n'ont sans doute jamais entendu les autres séquences ou ne les ont pas entendues aussi souvent et assez de fois pour les assimiler dans leur lexique.

Par contre, les productions élevées de liaisons interdites à toutes les étapes, et le fait que les productions de liaisons obligatoires, facultatives et interdites varient entre les temps 1, 2 et 3, suggèrent que puisque les sujets font des liaisons dans des contextes où ils n'ont probablement pas entendu de liaisons, ils sont en train de créer des catégories abstraites dans lesquelles ils généralisent, par analogie, les règles des liaisons obligatoires aux contextes de liaisons interdites.

Si le processus d'acquisition était arrêté et leur prononciation était fossilisée, on peut supposer que leurs productions ne changeraient pas beaucoup entre le temps 1, le temps 2 et le temps 3. Or, on observe des progrès et des changements dans tous les groupes, et ces progrès sont d'autant plus remarquables que la formation était très courte. Nous pensons donc que les sujets sont en cours d'acquisition et que des conditions favorisant l'apprentissage, en particulier un input enrichi avec une haute fréquence des structures cibles, de nombreuses répétitions d'items et structures permettrait aux apprenants de compléter leurs catégories abstraites et de mieux maîtriser le système des liaisons.

Un autre point à remarquer est que, contrairement aux francophones, les apprenants ne semblent pas connaître l'ordre hiérarchique des liaisons. Les priorités des francophones sont par ordre décroissant de faire d'abord une liaison obligatoire, puis un enchaînement et en dernier choix une liaison facultative. Or, le fait qu'au temps 3, les anglophones fassent plus de liaisons facultatives et d'enchaînements, mais moins de liaisons obligatoires que les francophones minoritaires, suggère qu'ils n'ont pas complètement l'intuition de ces priorités.



Ils peuvent par exemple faire des liaisons facultatives après *pas* ou après *est*, tandis qu'ils ne font pas les liaisons les plus importantes après les déterminants. Les anglophones ne traitent ni les catégories grammaticales, ni les consonnes de liaison comme les francophones et c'est surtout ici que se trouvent les plus grandes différences dans le traitement des liaisons.

Toutefois ils traitent la liaison de la même façon que les francophones en ce qui concerne la fréquence du mot1, du mot2 et la fréquence de cooccurrence du mot1 et du mot2, d'après ce qu'ils ont entendu. Plus les mots et les séquences sont fréquents dans l'input qu'ils ont reçu, plus les liaisons sont réalisées, cette tendance se retrouve avant la leçon sur les liaisons et après la leçon.

Etant donné qu'ils ont entendu les séquences beaucoup moins souvent que les francophones, ils traitent les séquences entendues comme les francophones qui ont peu entendu certaines séquences, notamment les séquences plus rares telles que dans les contextes [adj. + nom] et expressions figées. En effet, les francophones, eux aussi, ne font les liaisons dans ces contextes de façon systématique qu'avec les items les plus fréquents. Par exemple, les francophones majoritaires ont fait la liaison dans 100% des cas avec *moyen-âge*, mais dans seulement 58.3% des cas avec *accent aigu*, bien que la plupart des francophones de l'étude enseignent le français langue seconde / étrangère et aient sans doute entendu et prononcé cette séquence de nombreuses fois. Si l'on peut comprendre que les francophones qui n'enseignent pas le français puissent ne pas faire la liaison (et n'aient pas entendu et utilisé cette expression depuis longtemps), il est tout de même surprenant que ceux qui enseignent le français ne fassent pas cette liaison. Cela pourrait s'expliquer notamment par le manque de productivité de cette structure, car dans d'autres cas que les expressions figées on ne fait pas la liaison comme dans *accent étranger*.

De même, les francophones majoritaires ont fait les liaisons avec *léger ennui* dans 41.7% des cas, alors qu'ils les ont faites avec *premier amour* dans 91.7% des cas, bien qu'il

s'agisse pourtant des mêmes contextes phonétiques et syntaxiques. La fréquence lexicale joue un rôle encore plus déterminant pour les anglophones étant donné qu'ils n'ont pas encore une base de données suffisante pour avoir établi des représentations abstraites de catégories syntaxiques ou pour être capable de se fier aux facteurs phonétiques.

Ces résultats sont en accord avec le MBU. Si l'on observe l'acquisition du langage chez les enfants L1 (voir Ellis, 2002 et Tomasello, 2004), on remarque que les enfants apprennent d'abord des séquences isolées (*tokens*) sans être capable d'appliquer ce qu'ils savent de ces séquences à des séquences non entendues auparavant. À ce stade de leur acquisition, ils n'ont pas encore créé de représentations abstraites. Un peu plus tard, lorsqu'ils ont créé des représentations abstraites, les erreurs qu'ils font sont une généralisation des séquences les plus entendues (Dugua, 2005, 2006). Les adultes francophones en font de même : ils utilisent les règles générales qui consistent, par exemple, à ne pas faire la liaison dans le contexte [nom sing. + adj.] (*garçon intelligent*), c'est pourquoi ils ne font pas la liaison avec certaines expressions figées dans ce contexte (*accent aigu*), ou au contraire, font des liaisons interdites avec les H aspirés et font aussi des fausses liaisons. Ils utilisent leur banque de données pour traiter les informations selon un système probabiliste.

Les apprenants du français L2 ont eu un input limité qui ne leur a pas permis de construire des représentations abstraites. C'est la raison pour laquelle ils ne sont pas encore sensibles à la fréquence des structures grammaticales, mais sont sensibles à la fréquence lexicale, la fréquence des items. Ajouté à cela, leurs connaissances lexicales sont beaucoup plus limitées que celles des francophones, ce qui crée des difficultés évidentes notamment avec les expressions figées, ou avec les H aspirés qui sont des mots ou expressions assez rares. De plus, ils souffrent de l'interférence de leur langue maternelle, en particulier l'anglais.

En somme, les anglophones traitent les liaisons de la même façon que les francophones, le cerveau humain semble fonctionner de façon similaire en L1 et en L2, dans

le traitement des données en rapport avec la fréquence. Ils traitent des séquences selon leur banque de données et les items présents dans les catégories qui sont en cours de formation.

Ces items et ces catégories dépendent directement de leur fréquence. Or, comme la fréquence est réduite, les items le sont aussi, les items étant trop peu nombreux, ils ne suffisent pas à créer des catégories. Si ces catégories existent, elles sont très incomplètes. De plus, alors que les enfants francophones ont systématiquement entendu les noms ou adjectifs avec d'autres mots (déterminants), ils ont appris les noms avec leur déterminant et la liaison avec le déterminant, et même avant le nom isolé, ce qui crée d'ailleurs des erreurs telles que *le nouns* (Dugua. 2005).

Chevrot (2002), Chevrot et Fayol (2000) ainsi que Chabanal (2003) ont remarqué que les enfants apprennent les liaisons obligatoires avant les liaisons facultatives. En effet, les liaisons obligatoires sont plus fréquentes dans leur input, c'est pourquoi elles sont apprises en premier et sont le plus vite maîtrisées. Mais dans l'input des anglophones, les liaisons obligatoires ne sont guère plus fréquentes que certaines liaisons facultatives telles qu'avec la négation (*pas arrivé, pas agréable, pas encore...*), ou avec l'auxiliaire être + participe passé (*est arrivé*), ou avec être + adjectif ou adverbe (*c'est utile, c'est encore loin*). Ceci explique pourquoi ils n'ont pas de hiérarchie grammaticale comparable aux francophones et peuvent faire des liaisons facultatives tout en ne faisant pas certaines liaisons obligatoires. Par contre, ils ont acquis une hiérarchie lexicale, ce qu'on voit avec les items individuels et les résultats sur la fréquence.

L'effet de la fréquence se fait également ressentir avec le type d'instruction : les sujets ayant eu un input plus riche, un plus grand nombre d'heures d'instructions en français (ceux étant allés en école francophone ou ayant étudié le français intensif) ont de meilleurs résultats au début et à la fin, et ceux qui participent régulièrement à de nombreuses activités en français ont de meilleurs résultats que ceux qui ne participent à aucune activité ou à peu d'activités.

Enfin, l'enseignement de la prononciation et ici des liaisons semble permettre aux étudiants de progresser, et ceci est d'autant plus le cas dans le groupe 2, où les étudiants qui ont entendu et pratiqué le plus de répétitions ont fait plus de progrès que les autres groupes juste après les répétitions et même la semaine suivante. Une demi-heure d'enseignement a permis aux apprenants de faire des progrès dans le domaine des liaisons et des enchaînements.

Nous rejoignons ainsi l'opinion de Lauret (2007 : 39) selon laquelle « [...] la musique et les sons de la langue ne peuvent être adoptés que s'ils ont été largement entendus, écoutés précisément, commentés, distingués, 'apprivoisés' [...] ».

## **CHAPITRE V**

### **ANALYSES GOLDVARB ET DISCUSSION**

## Introduction

Nous pouvons maintenant répondre aux questions soulevées dans les hypothèses du chapitre 3 et confirmer les pourcentages relevés au chapitre 4. Dans ce chapitre, nous allons d'abord expliquer brièvement la démarche adoptée pour les analyses *GoldVarb* et les résultats obtenus avec ce logiciel, puis, nous discuterons des résultats généraux et de leurs implications.

Après avoir étudié les données à partir des pourcentages de réalisations (dont les principaux résultats sont au chapitre 4), nous avons effectué des analyses *GoldVarb* pour vérifier la signification statistique des tendances les plus pertinentes pour les hypothèses de notre étude. Nous avons d'abord procédé à une analyse individuelle des groupes de facteurs, dont nous n'incluons pas ici toutes les analyses, ceci étant surtout une analyse exploratoire visant à déterminer les facteurs ayant un effet significatif sur la liaison obligatoire, ou un effet qui montre une tendance, comme par exemple une relation linéaire avec le nombre d'années d'étude du français. Nous avons aussi écarté de cette étape de l'analyse les facteurs qui ne semblaient pas avoir un tel effet. Étant donné l'interdépendance entre certains groupes de facteurs (par exemple la fréquence du mot1 et la fréquence de cooccurrence), on ne peut pas combiner tous les groupes de facteurs codés dans la base de données.

Il a ensuite fallu recoder les groupes de facteurs quand leur nombre était trop élevé pour faire des généralisations, comme par exemple avec les contextes grammaticaux, ou lorsque les pourcentages de liaison semblaient proches (comme avec les différents déterminants). Nous avons ensuite effectué une analyse multivariationnelle sur les groupes de facteurs retenus. Enfin, nous avons effectué plusieurs analyses, notamment une analyse du progrès des apprenants entre les trois temps pour les liaisons obligatoires, d'abord pour tous les apprenants ensemble et ensuite pour chacun des 4 groupes. Les résultats obtenus confirment les hypothèses et renforcent les pourcentages déjà révélateurs.

## 1. Analyse GoldVarb

Nous allons ici analyser les éléments les plus représentatifs dans des tableaux et ferons quelques commentaires pour chaque tableau. Tous les groupes des facteurs dans ces tableaux ont un effet significatif sur la variable dépendante.

Nous voyons dans le tableau 5.1. que cette analyse confirme la signification des différences entre les trois groupes (francophones majoritaires, francophones minoritaires et anglophones) en ce qui concerne leur taux de liaisons obligatoires. La liaison est favorisée chez les deux groupes de francophones, mais elle est défavorisée chez les anglophones. Ceux-ci ne traitent donc pas les liaisons obligatoires de la même façon que les francophones.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .725				
<b>Groupe de participants:</b>				
Francophones majoritaires:	585	612	95.6%	0.891
Francophones minoritaires	349	408	85.5%	0.692
<b>Anglophones</b>	3760	5661	66.4%	0.429
écart				46
TOTAL	4694	6681	70.3%	

**Tableau 5.1. Taux de réalisation des liaisons obligatoires : Francophones majoritaires, francophones minoritaires, anglophones.**

L'analyse du taux de réalisation des liaisons facultatives (tableau 5.2.) confirme la signification des différences entre les trois groupes. La liaison facultative est favorisée chez les francophones majoritaires, mais elle est défavorisée chez les francophones minoritaires et les anglophones. Nous voyons ici que les francophones minoritaires ne traitent pas les liaisons facultatives comme les francophones majoritaires. Plusieurs liaisons facultatives se trouvant dans des structures syntaxiques plus rares (que les liaisons obligatoires), et n'étant pas systématiquement réalisées, les francophones minoritaires et les anglophones ne les ont sans doute pas assez entendues pour les avoir enregistrées dans leur système de classification. Par contre, ils réalisent les liaisons les plus fréquentes de cette catégorie (avec négation).

Même si les apprenants peuvent avoir été en contact avec un registre de langue plus recherché (et donc avoir entendu de nombreuses liaisons facultatives) que les francophones minoritaires qui ont entendu un français vernaculaire à la maison, à l'école, et surtout à l'université, leur exposition à la langue semble avoir trop limité pour avoir acquis toutes les liaisons obligatoires, et par extension celles qui sont encore moins fréquentes : les liaisons facultatives.

	Liaisons réalisées	N	%	Poids du facteur
Input: .255				
<b>Groupe de participants:</b>				
Francophones majoritaires:	76	204	37.3%	0.634
<b>Anglophones</b>	469	1887	24.9%	0.492
Francophones minoritaires	26	136	19.1%	0.409
écart				23
TOTAL	571	2227	25.6%	

**Tableau 5.2. Taux de réalisation des liaisons facultatives: Francophones majoritaires, francophones minoritaires, anglophones.**

L'analyse du taux de réalisation des liaisons interdites (tableau 5.3.) confirme la signification des différences entre les trois groupes. La non-liaison est favorisée chez les deux groupes de francophones, mais elle est défavorisée chez les anglophones.

Nous pouvons remarquer plusieurs éléments dans ce tableau. Tout d'abord, les productions de tous les groupes semblent plus éloignées des attentes du français standard (de ne pas faire la liaison), que dans les études précédentes. Dans 13.7% des cas, les francophones majoritaires ne les ont pas prononcées selon les normes, les francophones minoritaires dans 33% des cas, et les anglophones dans 44% des cas.

Rappelons que dans notre système de codage, seules les prononciations « standard » sont acceptées comme étant correctement réalisées. Le français franco-ontarien diffère du français standard dans plusieurs domaines. Par exemple, il est possible dans cette variété du français de prononcer le h des h aspirés. Cela explique en partie un taux plus bas de



réalisations standard de liaisons interdites (67%). Notons également que certains francophones majoritaires ont, eux aussi, prononcé le h (peut-être à cause de l'influence de l'anglais ?).

Quant aux anglophones, d'une part, ils ont parfois prononcé le h du mot2, d'autre part, ils ont parfois prononcé la consonne finale du mot1, et enfin ils ont aussi fait la liaison. C'est la raison pour laquelle nous avons choisi de parler des cas correctement réalisés, plutôt que de liaisons effectuées, car d'autres cas peuvent rentrer en compte pour ne pas prononcer une séquence selon les règles du français standard (tels que la prononciation de h, prononciation de consonne finale muette, etc.).

	<b>Liaisons non réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .602				
<b>Groupe de participants:</b>				
Francophones majoritaires:	145	168	86.3%	0.806
Francophones minoritaires	75	112	67.0%	0.572
<b>Anglophones</b>	870	1554	56.0%	0.456
écart				34
TOTAL	1090	1834	59.4%	

**Tableau 5.3. Taux de réalisation des liaisons interdites Francophones majoritaires, francophones minoritaires, anglophones.**

L'analyse du taux de réalisation des enchaînements (tableau 5.4.) confirme la signification des différences entre les trois groupes. L'enchaînement est favorisé chez les francophones majoritaires, mais il est défavorisé chez francophones minoritaires et chez les anglophones. Il est intéressant de remarquer ici que les apprenants ont plus tendance à faire les enchaînements que les francophones minoritaires. Les Franco-ontariens pourraient avoir des « priorités différentes » des apprenants et se concentrer sur les liaisons obligatoires, tandis que les anglophones seraient moins « sélectifs » quant à la hiérarchie des situations liantes.

	<b>Enchaînements réalisés</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .614				
<b>Groupe de participants:</b>				
Francophones majoritaires:	122	156	78.2%	0.692
Francophones minoritaires	57	104	54.8%	0.432
<b>Anglophones</b>	863	1443	59.8%	0.483
écart				21
TOTAL	1042	1703	61.2%	

**Tableau 5.4. Taux de réalisation des enchaînements: Francophones majoritaires, francophones minoritaires, anglophones.**

En ce qui concerne le rôle de la fréquence des mots1 (tableau 5.5.) sur la réalisation de la liaison chez les francophones, la liaison est favorisée pour les mots les plus fréquents et défavorisée pour les moyennement fréquents et les moins fréquents. Les liaisons sont réalisées à 99.2% par tous les francophones avec les mots1 les plus fréquents, à 91.2% avec les mots moyennement fréquents, et à 80.8% avec les mots les moins fréquents.

Nous avons recodé les groupes de facteurs en trois tranches : mots les plus fréquents (3), moyennement fréquents (2) et moins fréquents (1) avec la banque de données de *Lexique*. Il apparaît ainsi que, plus les mots sont fréquents, plus les francophones réalisent les liaisons.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .952				
<b>Fréquence mot1 Lexique:</b>				
3 (i j k): plus fréquent	357	360	99.2%	0.857
2 (d e f g h) moyennement fréquent	383	420	91.2%	0.343
1 (a b c) moins fréquent	194	240	80.8%	0.175
écart				68
TOTAL	934	1020	91.6%	

**Tableau 5.5. Liaisons obligatoires, francophones. Fréquence Lexique mot1**

En ce qui concerne le rôle de la fréquence des mots1 (tableau 5.6.) sur la réalisation de liaisons obligatoires chez les anglophones, la liaison est favorisée pour les mots les plus

fréquents et défavorisée pour les moyennement fréquents et les moins fréquents. Il est intéressant de remarquer que c'est ici une tendance similaire pour les francophones et les anglophones. Les anglophones utilisent donc la fréquence des mots1 pour traiter les liaisons.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .673				
<b>Fréquence mot1 Lexique:</b>				
3 (i j k): plus fréquent	1576	1998	78.9%	0.644
2 (d e f g h) moyennement fréquent	1516	2331	65.0%	0.475
1 (a b c) moins fréquent	668	1332	50.2%	0.329
écart				31
TOTAL	3760	5661	66.4%	

**Tableau 5.6. Liaisons obligatoires, anglophones. Fréquence *Lexique* mot1**

En ce qui concerne le rôle de la fréquence des mots2 (tableau 5.7.) sur la réalisation de la liaison chez les francophones, la liaison est favorisée pour les mots les plus fréquents et moyennement fréquents et elle est défavorisée pour les mots les moins fréquents. Plus les mots2 sont fréquents, plus les liaisons sont réalisées. Plus les mots2 sont rares, moins les liaisons sont réalisées. La fréquence des mots2 est donc un indice fiable dans la production de liaisons des francophones.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .942				
<b>Fréquence mot2 Lexique:</b>				
3 (i j k): plus fréquent	79	80	98.8%	0.830
2 (d e f g h) moyennement fréquent	429	440	97.5%	0.707
1 (a b c) moins fréquent	426	500	85.2%	0.263
écart				57
TOTAL	934	1020	91.6%	

**Tableau 5.7. Liaisons obligatoires, francophones. Fréquence *Lexique* mot2**

En ce qui concerne le rôle de la fréquence des mots2 (tableau 5.8.) sur la réalisation de liaisons obligatoires chez les anglophones, la liaison est favorisée pour les mots les plus fréquents et moyennement fréquents et défavorisée dans les mots les moins fréquents. Les anglophones utilisent donc la fréquence des mots2 pour traiter les liaisons.

Toutefois, les anglophones ont tendance à réaliser davantage la liaison avec les mots moyennement fréquents. Ces résultats sont légèrement différents des résultats annoncés dans les chapitres précédents parce qu'il y a eu un recodage en incluant différents groupes (ici d, e, f, g, h) dans la catégorie « moyennement fréquente ».

Ainsi, les apprenants peuvent réaliser davantage de liaisons avec des mots moyennement fréquents qu'avec des mots plus fréquents. Cela pourrait être dû au fait que la notion de fréquence dépend pour les apprenants de la fréquence des mots dans leur input.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .674				
<b>Fréquence mot2 Lexique:</b>				
3 (i j k): plus fréquents	330	444	74.3%	0.583
2 (d e f g h) moyennement fréquents	1901	2442	77.8%	0.629
1 (a b c) moins fréquents	1529	2775	55.1%	0.373
écart				26
TOTAL	3760	5661	66.4%	

**Tableau 5.8. Liaisons obligatoires anglophones. Fréquence *Lexique* mot2**

Pour ce qui est du rôle de la fréquence de cooccurrence des mots1 et mots2 dans *Lexique* dans la réalisation des liaisons obligatoires chez les francophones (tableau 5.9.), la liaison est favorisée dans les séquences au taux de cooccurrence le plus élevé et elle est défavorisée pour les deux autres catégories. Globalement, on voit que plus le taux de cooccurrence est élevé, plus les liaisons sont réalisées pour tous les groupes de francophones.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .943				
<b>Fréquence cooccurrence Lexique:</b>				
3 (i j k): plus fréquent	277	280	98.9%	0.848
2 (d e f g h) moyennement fréquent	245	260	94.2%	0.498
1 (a b c) moins fréquent	412	480	85.8%	0.269
écart				55
TOTAL	934	1020	91.6%	

**Tableau 5.9. Liaisons obligatoires, francophones. Fréquence *Lexique* cooccurrence**

On observe la même tendance chez les anglophones (tableau 5.10.) : plus le taux de cooccurrence est élevé, plus les anglophones ont tendance à faire la liaison, mais dans en plus faible proportions que pour les francophones.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .676				
<b>Fréquence cooccurrence Lexique:</b>				
3 (i j k): plus fréquent	1241	1554	79.9%	0.655
2 (d e f g h) moyennement fréquent	1093	1443	75.7%	0.599
1 (a b c) moins fréquent	1426	2664	53.5%	0.356
écart				30
TOTAL	3760	5661	66.4%	

**Tableau 5.10. Liaisons obligatoires, anglophones. Fréquence *Lexique* cooccurrence**

En ce qui concerne le rôle des facteurs linguistiques dans la réalisation de liaisons obligatoires pour les trois groupes (tableau 5.11.), nous avons classé les groupes de facteurs selon l'ordre d'importance de leur effet sur la liaison (écart type) ; cela confirme la tendance selon laquelle les facteurs linguistiques ont un effet plus important que les facteurs sociaux.

Nous avons recodé les facteurs grammaticaux dans les différents groupes, selon des caractéristiques communes qui semblent aussi refléter des effets proches (d'après les

pourcentages) sur les liaisons obligatoires.

On observe donc le rôle des facteurs phonétiques, syntaxiques et fréquentiels dans le traitement de la liaison. Cependant, les données ci-dessous comprennent les données des francophones majoritaires, minoritaires et le groupe des apprenants. Il s'agit alors de déterminer avec plus de précision, les éléments ci-dessous pour chaque groupe.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: 0.751				
<b>Groupe de participants:</b>				
francophones majoritaires:	585	612	95.6%	0.914
francophones minoritaires	349	408	85.5%	0.720
apprenants	3760	5661	66.4%	0.420
écart				49
<b>Consonne de liaison</b>				
z	2756	3668	75.1%	0.590
n	1196	1572	76.1%	0.530
t	474	917	51.7%	0.285
r	199	393	50.6%	0.259
v	69	131	52.7%	0.110
écart				48
<b>Fréquence de cooccurrence</b>				
2 (d e f g h) moyennement fréquent	1338	1703	78.6%	0.720
3 (i j k): plus fréquent	1518	1834	82.8%	0.543
1 (a b c) moins fréquent	1838	3144	58.5%	0.352
écart				37
<b>Contexte grammatical</b>				
[det monosyll. + nom]	2537	3144	80.7%	0.559
[prép (+ art.) + nom]	406	524	77.5%	0.544
[adj + nom; nom + adj]	1495	2489	60.1%	0.460
[det 2 syllabes + nom]	256	524	48.9%	0.302
écart				36
<b>TOTAL</b>	<b>4694</b>	<b>6681</b>	<b>70.3%</b>	

**Tableau 5.11. Facteurs linguistiques et liaisons obligatoires pour les trois groupes**

Vérifions plus précisément le rôle des facteurs phonétiques, syntaxiques et fréquentiels dans le traitement de la liaison chez les apprenants.

Pour analyser le rôle des facteurs linguistiques et individuels/sociaux chez les apprenants sur les productions de liaisons obligatoires (tableau 5.12.), nous avons également classé les résultats selon l'ordre d'importance de leur effet sur la liaison (écart).

Les facteurs linguistiques ont le poids le plus fort. Les anglophones ont tendance à faire les liaisons avec les consonnes /z/ et /n/, avec les séquences aux mots les plus fréquents et au taux de cooccurrence le plus élevé, et dans certains contextes grammaticaux. Les personnes participant à plus de 3 activités en français font aussi plus de liaisons obligatoires et celles qui ont étudié le français plus de 9 années ont tendance à réaliser un plus grand nombre de liaisons obligatoires.

Par contre, le facteur sexe est dans le sens contraire de celui observé pour les francophones. En effet, ce facteur était corrélé à d'autres facteurs chez les francophones, en particulier les conditions d'apprentissage du français (les sujets franco-ontariens étaient des femmes), à l'âge et au niveau d'étude. Par contre, si les conditions d'apprentissage sont globalement similaires, ce qui est le cas ici, alors le facteur sexe n'est pas significatif.

Il ressort de ceci que les anglophones utilisent des informations phonétiques et fréquentielles, et dans une moindre mesure des informations syntaxiques, pour traiter les informations. Toutefois, la façon dont ils le font est différente des francophones sur certains points.

Globalement, on peut estimer que les tendances sont les mêmes que les francophones : les apprenants ont tendance à réaliser les liaisons avec les consonnes de liaison les plus fréquentes /z/ et /n/, mais ils les réalisent en bien plus faible proportion que les francophones. D'autre part, les apprenants ont par exemple tendance à ne pas faire la liaison avec le /v/ de *neuf*, alors que cette liaison est systématique chez les francophones. Tout comme les francophones, ils font moins la liaison avec le /R/ qui est, en effet, une liaison plus rare que les autres. Toutefois, les tendances sont beaucoup plus basses.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: 0.691				
<b>Consonne de liaison</b>				
z	2211	3108	71.1%	0.585
n	963	1332	72.3%	0.525
r	165	333	49.5%	0.299
t	372	777	47.9%	0.297
v	49	111	44.1%	0.098
écart				57
<b>Fréquence de co-occurrence</b>				
2 (d e f g h) moyennement fréquent	1093	1443	75.7%	0.719
3 (i j k): plus fréquent	1241	1554	79.9%	0.550
1 (a b c) moins fréquent	1426	2664	53.5%	0.349
écart				43
<b>Contexte grammatical</b>				
[det monosyll. + nom]	2060	2664	77.3%	0.558
[prép (+ art.) + nom]	327	444	73.6%	0.544
[adj + nom; nom + adj]	1177	2109	55.8%	0.463
[det 2 syllabes + nom]	196	444	44.1%	0.294
écart				27
<b>Activités en français</b>				
3 ou plus d'activités	1545	2142	72.1%	0.560
1 - 2 activités	1731	2601	66.6%	0.499
aucune	484	918	52.7%	0.366
écart				19
<b>Années de français</b>				
9 ans ou plus	949	1224	77.5%	0.631
1 à 8 ans	2811	4437	63.4%	0.463
écart				17
<b>Sexe</b>				
femmes	3222	4743	67.9%	0.524
hommes	538	918	58.6%	0.376
écart				14
<b>TOTAL</b>	<b>3760</b>	<b>5661</b>	<b>66.4%</b>	

**Tableau 5.12. Liaisons obligatoires, facteurs linguistiques et individuels/sociaux chez les anglophones**

Si l'on considère les réalisations de liaisons obligatoires de tous les apprenants aux trois temps (tableau 5.13.), nous avons la confirmation de la signification statistique des



tendances aux progrès entre les temps 1, 2 et 3., et donc la confirmation du rôle de la leçon de phonétique.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .665				
<b>Temps de lecture:</b>				
Temps 1	1145	1887	60.7%	0.437
Temps 2	1304	1887	69.1%	0.529
Temps 3	1311	1887	69.5%	0.534
écart				11
TOTAL	3760	5661	66.4%	

**Tableau 5.13. Liaisons obligatoires, tous apprenants. Trois Temps.**

Si l'on considère les réalisations de liaisons obligatoires des apprenants du groupe 1 aux trois temps (tableau 5.14.), il n'y a pas de différences significatives entre les trois temps.

Comme nous l'avons vu dans la méthodologie, le groupe 1 est celui qui a reçu le plus d'explications et le moins de corrections. Il semble donc que les explications soient, parmi tous les éléments testés dans la leçon de phonétique, soient le facteur qui joue le rôle le moins important dans les progrès des étudiants. Les corrections sont peut-être aussi plus importantes que les explications.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: .673				
<b>Temps de l'enregistrement</b>				
Temps 1	325	510	63.7%	
Temps 2	359	510	70.4%	n.s.
Temps 3	346	510	67.8%	
écart				
TOTAL	1030	1530	67.3%	

**Tableau 5.14. Liaisons obligatoires, anglophones, Groupe 1. Trois Temps.**

En analysant les réalisations de liaisons obligatoires des apprenants du groupe 2, aux trois temps (tableau 5.15.), on observe un effet significatif de l'enseignement, car il y a eu des progrès entre les 3 temps, et ce, juste après la leçon, et une semaine plus tard. Rappelons que le groupe 2 est celui qui avait bénéficié des répétitions au moment du test dans la leçon de phonétique.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: 0.611				
<b>Temps de l'enregistrement</b>				
Temps 1	268	510	52.5%	0.414
Temps 2	328	510	64.3%	0.535
Temps 3	336	510	65.9%	0.552
écart				14
TOTAL	932	1530	60.9%	

**Tableau 5.15. Liaisons obligatoires, anglophones, Groupe 2. Trois Temps.**

Pour ce qui est des progrès du groupe 3 (tableau 5.16.), on remarque des progrès entre les trois temps, mais dans une moindre mesure que pour le groupe 2. Ce groupe avait aussi eu des répétitions, mais un peu moins que le groupe 2.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input: 0.690				
<b>Temps de l'enregistrement</b>				
Temps 3	549	765	71.8%	0.533
Temps 2	544	765	71.1%	0.525
Temps 1	487	765	63.7%	0.441
écart				9
TOTAL	1580	2295	68.8%	

**Tableau 5.16. Liaisons obligatoires, anglophones, Groupe 3. Trois Temps.**

Quant au groupe 4 (tableau 5.17.) qui n'avait pas eu de formation phonétique au moment du test, il n'y a pas de différence significative entre les trois temps pour les taux de liaison obligatoires.

	<b>Liaisons réalisées</b>	<b>N</b>	<b>%</b>	<b>Poids du facteur</b>
Input:				
<b>Temps de l'enregistrement</b>				
Temps 1	65	102	63.7%	
Temps 2	73	102	71.6%	n.s.
Temps 3	80	102	78.4%	
écart				
TOTAL	218	306	71.2%	

**Tableau 5.17. Liaisons obligatoires, anglophones, Groupe 4. Trois Temps.**

## 2. Discussion sur les résultats obtenus

Les francophones majoritaires, les francophones minoritaires et les apprenants L2 utilisent la fréquence relative aux mots1, aux mots2 et la fréquence de cooccurrence des mots1 et des mots2 pour traiter les informations et produire les liaisons. En fait, c'est cette fréquence qui semble jouer le rôle le plus important pour les trois groupes. Le mot1 indique en partie la fréquence syntaxique et en partie la fréquence lexicale, tandis que le mot2 représente uniquement la fréquence lexicale (la fréquence des items). La fréquence des mots1 joue un rôle aussi important que la fréquence des mots2 dans le traitement des liaisons pour tous les groupes, et ces informations sont renforcées par la fréquence de cooccurrence, c'est-à-dire que s'opère un cumul des facteurs de fréquence lexicale.

Si on analyse la réalisation des liaisons obligatoires par groupe syntaxique, on remarque que le facteur syntaxique joue un rôle évident pour les francophones et un certain rôle pour les apprenants (tableau 5.12.). Les anglophones réalisent en plus grandes proportions les structures les plus fréquentes telles que [dét. monosyllabique + nom] et réalisent en plus faible proportion les structures syntaxiques moins fréquentes et moins productives de type [adj. + nom] ou [dét. 2 syllabes + nom]. Ces tendances sont similaires à celles des francophones.

Toutes les catégories ont été réalisées à plus de 97% par tous les francophones (majoritaires et minoritaires), sauf les catégories [adj. qual. + nom] (86.6%), [nom + adj.] notamment dans les expressions figées (68.3%) et dans la catégorie [adj. ord. + nom] (52.5%). Les liaisons sont réalisées en plus faibles proportions dans les catégories syntaxiques étant les moins productives (regroupant un nombre plus restreint d'items).

Les expressions figées sont réalisées à 87.5% pour les francophones majoritaires, et à 72.91% par les francophones minoritaires, à 62.6% par les apprenants L2 au temps 1 et à 72.1.% au temps 3. Il est intéressant de remarquer que les productions de liaisons dans les expressions figées des apprenants avoisinent au temps 3 celles des francophones minoritaires. Ceci est peut-être dû à la leçon de phonétique dans laquelle les expressions figées en question étaient présentes. De plus, les apprenants L2 ont sans doute entendu la séquence « accent aigu » plus souvent que les francophones, puisque c'est fréquemment employé en salle de classe en L2.

On observe un effet cumulé de la fréquence et de la productivité des structures avec la fréquence des items. Si des items sont peu fréquents et appartiennent à des structures peu fréquentes, alors les productions sont faibles telles que *léger ennui* réalisé à 30% par tous les francophones. Si un item est très fréquent dans une structure peu productive, la fréquence de l'item compense le manque de productivité de la structure comme avec *États-Unis* dont la liaison a été réalisée à 100% par tous les francophones. On remarque la même tendance avec les anglophones, qui réalisent la liaison à 18.9% avec *léger ennui* au temps 3, mais à 97.3% avec *États-Unis* au temps 3.

En revanche, on ne peut pas en déduire que la liaison avec *États-Unis* a été réalisée dans une telle proportion (à 97.3%), simplement parce que les apprenants savent qu'un trait d'union signifie qu'il s'agit d'un seul mot (composé de deux mots). En effet, ils n'ont produit la liaison (correcte), au temps 3, que dans 75.7% avec *sous-entendu* et 70.3% avec *faux-airs*.

Si l'on considère l'ensemble des contextes de sandhi, les francophones majoritaires ont produit 95.6% de liaisons obligatoires, 78.2% d'enchaînements, 37.3% de liaisons facultatives et ont produit correctement 86.3% de liaisons interdites (donc non réalisées, soit 13.7% de liaisons interdites produites). Les francophones minoritaires ont produit 85.5% de liaisons obligatoires, 54.8% d'enchaînements, 19.1% de liaisons facultatives et 67% de liaisons interdites (soit 33% produites ou réalisées différemment des normes orthoépiques).

Les apprenants L2 ont produit 60.7% de liaisons obligatoires au temps 1 et 69.5% au temps 3, 55.1% d'enchaînements au temps 1 et 60.9% au temps 3, 22.4% de liaisons facultatives au temps 1 et 26.6% au temps 3. Ils ont donc fait des progrès dans tous les domaines, et c'est avec les mots1 et mots2 les moins fréquents qu'ils ont fait le plus de progrès.

Bien que la leçon de phonétique n'ait abordé que très brièvement les liaisons interdites et se soit concentrée sur les liaisons obligatoires, les apprenants ont produit un taux plus élevé de liaisons facultatives et d'enchaînements au temps 2 et au temps 3 (qu'au temps 1), alors que la leçon ne portait pas sur ces domaines. Il faut préciser que dans les liaisons facultatives se trouvaient certains items fréquents tels que les auxiliaires (*sont allés, ont attendu, ont écrit, ont eu*), le verbe être (*est un*), et la négation *pas* (*pas allés*). Il y avait aussi des items très fréquents dans les enchaînements (*cet homme, il a, elle a, il aura*).

C'est avec les liaisons obligatoires que les différences sont les plus visibles entre les francophones et les apprenants. Les apprenants ont produit 60.7% de liaisons obligatoires au temps 1 et 69.5% au temps 3, ce qui constitue une énorme différence avec les francophones majoritaires (95.6%). De plus, les 60.7% de liaisons obligatoires des anglophones de notre étude sont également très éloignés des 91.1% de l'étude de Thomas (2002), des 95.36 % de l'étude de Mastromonaco (2000) ou des 95% de l'étude de Howard (2005).

Il est vrai que les sujets de ces études étaient en 2<sup>e</sup> ou 3<sup>e</sup> année d'étude, alors que les sujets de notre étude n'étaient qu'en 1<sup>e</sup> année. On pourrait penser que les sujets de ces études ont obtenu de meilleurs résultats parce qu'ils avaient un niveau plus avancé. Toutefois, nous pensons que les taux de réalisation des liaisons beaucoup plus bas de nos sujets ne sont pas dus à une véritable différence de niveau, mais plutôt aux méthodes utilisées pour tester les apprenants (dans le test lui-même et le codage). En fait, les pourcentages de Thomas (2002 et 2004) sur les non-enchaînements des liaisons (8.5%) sont incroyablement proches des nôtres (8.85%). Cela suggère que les sujets de son étude ont un niveau très proche des nôtres.

Une différence entre l'étude de Thomas et la nôtre est que Thomas considère que, puisque la consonne finale a été prononcée, c'est une forme de liaison, donc faisant partie des liaisons prononcées (d'où des chiffres plus élevés que les autres). Par contre, nous avons considéré ces cas comme ne faisant pas partie des liaisons, car celles-ci ne sont pas correctement réalisées. À notre avis, ce ne sont pas des liaisons, mais simplement des consonnes finales prononcées, consonnes finales qui seraient prononcées même si le mot suivant commençait avec une consonne (comme nous l'avons remarqué lors du codage des enregistrements).

Une autre catégorie qui est réalisée de façon similaire chez les sujets de Thomas (2002 et 2004) et les nôtres est la catégorie [adj. + nom] réalisée à 49.7% dans l'étude de Thomas (2002) et à 48.8% dans notre étude alors que les francophones majoritaires de notre étude réalisent les liaisons avec les adjectifs qualificatifs à 94.8%.

Si on analyse les productions de liaisons d'après les catégories syntaxiques, on peut considérer que les apprenants n'ont pas encore développé des catégories syntaxiques abstraites assez complètes pour pouvoir réaliser les liaisons de façon systématique d'après ce critère, y compris dans les structures les plus fréquentes de type [art. déf. + nom], catégorie pour laquelle ils produisent 67.6% de liaisons obligatoires au temps 1 et 74.3% au temps 3.

Un phénomène intéressant est que l'ordre d'acquisition, classé par ordre décroissant d'après le pourcentage de réalisation des liaisons au temps 3 (informations du tableau 4.48.) est assez différent de celui auquel on s'attendrait, si l'on considère l'ordre d'acquisition des enfants francophones : [adj. pos. + nom] : 85,7 % au temps 1 et 91,5% au temps 3. [adj. num. + chif.] : 64,9% au temps 1 et 87,8 % au temps 3. [adj. dém. + nom] : 70,3% au temps 1 et 81,1% au temps 3. [prép. + art. + nom] : 69,4% au temps 1 et 76,6% au temps 3. [prép. + nom] : 73% au temps 1 et 75,7% au temps 3. [adj. num. + nom] : 68,7% au temps 1 et 75,3% au temps 3. [art. déf. + nom] : 67,6% au temps 1 et 74,3% au temps 3. [art. ind. + nom] : 61,6% au temps 1 et 73,5% au temps 3. [nom + adj.] : 50,5% au temps 1 et 66,7% au temps 3. [adj. ind. + nom] : 50% au temps 1 et 58,1% au temps 3. [adj. qual. + nom] : 48,8% au temps 1 et 57,1% au temps 3. [adj. ord. + nom] : 28,4% au temps 1 et 33,8% au temps 3.

On voit ici que les liaisons avec les articles définis et indéfinis sont réalisées en beaucoup plus faible quantité qu'avec les adjectifs possessifs, numéraux et démonstratifs par exemple. Les apprenants ont peut-être appris à l'école des structures préfabriquées du type : *mes affaires, mon avis*, etc. dans lesquelles se trouvent ces structures grammaticales. Ces résultats suggèrent que les apprenants ne maîtrisent pas complètement les structures syntaxiques les plus fréquentes réalisées à 100% par tous les francophones, car si c'était le cas, les liaisons seraient produites de façon systématique dans ces catégories.

Les catégories syntaxiques qui semblent être le mieux maîtrisées sont les adjectifs possessifs (91.5%) et numéraux (87.8%). Rappelons que la catégorie [adj. num. + chif.] réalisée à 87,8 % au temps 3 représente la prononciation du /t/ de *vingt-trois* et de *vingt-cinq*. Cela veut dire que dans 14.3% des cas, les apprenants ayant étudié le français au moins 6 ans ne prononcent pas (correctement) le /t/ de *vingt* au temps 1 et 12.2% au temps 3.

Les séquences contenant des chiffres (*deux amis, dix ans, six heures, neuf heures, cent ans...*) sont pourtant des séquences fréquentes, mais dans 31.3% des cas, les apprenants ne

réalisent pas correctement ces liaisons au temps 1 et 24.7% au temps 3. Soit ils ne prononcent pas correctement la consonne de liaison et la prononcent /s/ au lieu de /z/, /f/ au lieu de /v/, soit ils n'enchaînent pas la liaison. Dans 32.4% des cas, les liaisons avec les articles définis ne sont pas réalisées au temps 1 et 25.7% ne sont pas réalisées au temps 3. De même, dans 38.4% des cas, les liaisons n'ont pas été réalisées correctement avec les articles indéfinis et 26.5% au temps 3. Dans un grand nombre de cas, lorsque la liaison n'a pas été réalisée correctement, c'était parce que les consonnes n'ont pas été enchaînées (*un accent, un hôpital, un oiseau, des histoires ...*), en particulier dans les séquences contenant l'article *un*.

Ces pourcentages de réalisation peu élevés étaient attendus pour différentes raisons. D'une part, seules les liaisons « correctement réalisées » avec la bonne consonne de liaison et enchaînée ont été considérées comme telles, alors que dans les études de Thomas (2002) ou Mastromonaco (2000), le système de codage était différent. D'autre part, les éléments testés représentent de plus grandes difficultés pour les apprenants, tels que les mots commençant par un h ou une semi-consonne (*un oiseau*).

Le taux de réalisation peu élevé de liaisons des apprenants L2 peut d'abord paraître un peu décourageant. Produire 52.3% de liaisons obligatoires après au moins 6 années d'étude du français, dont 67,6% avec les articles définis et 61,6% avec les articles indéfinis au temps 1 est un peu décevant.

Par contre, les progrès des étudiants après la leçon de phonétique sont encourageants. Ils indiquent que leur prononciation n'est pas complètement fossilisée et qu'ils peuvent encore progresser, ce qu'ils ont réussi à faire en un temps très restreint. En effet, si leur interlangue était complètement fossilisée, ils ne progresseraient plus. Au contraire, ils ont fait des progrès remarquables dans presque toutes les catégories (sauf les liaisons interdites), y compris celles qui n'ont pas vraiment été enseignées : les liaisons facultatives et les enchaînements. Quel que soit le type d'enseignement reçu à l'école primaire et secondaire, les



étudiants issus de l'immersion, du français intensif, ou du français cadre ont tous fait des progrès. La moyenne de tous les groupes était 60.7% au temps 1 et 69.5% au temps 3, soit des progrès de presque 10% après seulement 30mn de cours, et ce une semaine après le cours, dans le post-test.

Un autre phénomène qui peut être considéré comme un signe de développement langagier est le taux assez élevé de réalisation de liaisons interdites. S'ils produisent des liaisons dans des séquences dans lesquelles ils n'ont probablement jamais entendu de liaisons cela indique qu'ils traitent ces exceptions par analogie aux structures plus fréquentes (les liaisons obligatoires), et qu'ils sont en train de développer un système de catégories abstraites, sauf dans les cas où les erreurs consistent à prononcer les consonnes finales des mots<sup>1</sup>.

Par contre, leurs progrès semblent partiellement limités par la langue maternelle qui semble jouer un rôle dans la production de liaisons, tout comme dans le non-enchaînement de liaisons. Les locuteurs d'autres langues que le chinois et l'anglais produisent 68.6% de liaisons obligatoires au temps 1 et 82.4% au temps 3 et ce sont les productions les plus élevées. Les sujets bilingues produisent 66.1% de liaisons au temps 1 et 76.8% au temps 3. Le locuteur chinois produit 60.8% de liaisons au temps 1 et 74.8% au temps 3. Enfin, les anglophones produisent 58.7% de liaisons au temps 1 et 66.4% au temps 3.

On remarque aussi que les locuteurs anglophones monolingues produisent le plus de liaisons non-enchaînées au temps 3 (9.80% au temps 1 et 7.41% au temps 3), tandis que les sujets d'autres langues maternelles en produisent 10.78% au temps 1 et 6.86% au temps 3, les locuteurs bilingues en produisent 5.88% au temps 1 et 4.76% au temps 3, et enfin le locuteur chinois, en produit 0% au temps 1 et 0% au temps 3.

On observe aussi des différences dans le traitement des consonnes de liaison selon les langues maternelles. Par exemple les liaisons obligatoires sont réalisées avec /R/ à 41.98% par les anglophones, à 47.62% par les bilingues, à 66.67% par le locuteur chinois et à 83.33%

par les germanophones et arabophones. Autrement dit, selon la langue maternelle des locuteurs, la production selon la consonne de liaison passe du simple au double.

Il faudrait analyser le système syllabique de différentes langues pour en venir à des conclusions définitives, ce qui n'est pas le sujet de notre thèse. Nous nous restreindrons à la conclusion très générale selon laquelle la langue maternelle semble avoir un rôle dans la production de liaisons enchaînées et de liaisons non- enchaînées.

Puisque l'input est un facteur clé dans la production de liaisons, le type d'input reçu durant l'apprentissage du français est important. Plus les sujets ont reçu un input riche, plus leur production est élevée. Le sujet ayant fréquenté une école francophone quelques années à produit 82.4% de liaisons obligatoires, ce qui se rapproche des 85.5% des francophones minoritaires et en particulier des Africains (84.8%). Les sujets ayant étudié le français intensif ont de meilleurs résultats que les autres (77.1% au temps 3).

Si les apprenants n'ont ni bénéficié de l'input leur permettant d'apprendre certaines structures naturellement, ni bénéficié d'instruction formelle leur permettant d'apprendre ces structures en classe, ils ont eu recours à leur langue maternelle (l'anglais) pour bâtir le système phonétique de la langue cible, et leur prononciation s'est (partiellement) fossilisée. C'est la raison pour laquelle on retrouve certains traits de l'anglais, notamment de syllabation, avec une tendance à prononcer les consonnes finales sans les enchaîner.

Par contre, une autre question se pose en ce qui concerne les étudiants venant d'immersion. Certes, ils ont une certaine facilité d'élocution, une grande confiance en eux et ont une excellente compréhension écrite et orale de la L2. Ils ont pratiqué le français beaucoup plus que ceux du français de base et ils ont pourtant produit 60.5% de liaisons obligatoires, alors que ceux venant du français de base en ont produit 62.4% en ayant étudié le français seulement 3-4 heures par semaine. De plus, on pourrait penser que puisque les étudiants venant d'immersion ont reçu un input plus riche, ils devraient faire plus de progrès

que les étudiants issus du français de base.

En effet, en admettant que la qualité du français oral de certains enseignants et des élèves ne soit pas aussi bonne qu'on le souhaiterait, et que ce français fautif ait fait parti de leur input, les étudiants venant d'immersion ont quand même été en contact avec un français authentique beaucoup plus que ceux du français de base, en regardant des émissions de télévision, en écoutant des chansons dont les locuteurs sont généralement des francophones (majoritaires). De plus, bon nombre d'enseignants sont francophones et d'autres enseignants non-natifs parlent le français à un niveau quasi-francophone.

Or, au temps 3, les étudiants d'immersion en produisent 70.3% et ceux du français de base 70.1%, soit globalement les mêmes résultats. Swain (1985) observe que les productions écrites et orales des élève d'immersion sont moins bonnes que leur compréhension écrite et orale. Cela pourrait impliquer que l'exposition à la langue (en particulier en réception) pourrait ne pas être suffisante pour acquérir certaines structures syllabiques ou phonétiques.

Tout comme les étudiants venant d'immersion produisent moins de liaisons obligatoires (60.5%) que ceux venant du français cadre (au temps 1), le fait que les sujets de notre étude ayant appris le français pendant 13-14 ans ne produisent que 53.7% de liaisons obligatoires au temps 1 suggère qu'il pourrait y avoir eu une fossilisation de certains aspects phonétiques et que, là aussi, l'exposition à la langue pourrait être insuffisante.

Selon Thomas (1998), alors que les francophones apprennent la liaison et tout son système naturellement, ce n'est pas le cas pour les élèves d'immersion qui, après six années d'immersion, n'ont pas appris la liaison en écoutant et répétant. Ils ne connaissent pas les phénomènes de liaison, car la liaison est rarement enseignée, et elle est souvent mal pratiquée dans leur entourage, car elle n'entrave que rarement la communication.

Dans notre étude, les anglophones prononcent les consonnes de liaisons sans les enchaîner dans 8.85% des cas au temps 1 et continuent à ne pas les enchaîner dans 6.68% des

cas au temps 3, alors que ce point a été abordé en classe et que les étudiants ont été corrigés sur ces erreurs-là.. Ceci suggère aussi une certaine fossilisation. L'écoute, la répétition, les explications et les corrections (limitées à ½ heure), ne suffisent pas à changer cette habitude.

Ceci a pour cause l'influence du système phonétique de l'anglais qui consiste à prononcer les consonnes finales. Les différences entre le temps 1 et le temps 2 indiquent qu'il est possible de faire des progrès (6.68% contre 8.85%), mais que ces progrès sont lents et limités dus à la fossilisation de l'interlangue. Car, comme le dit Lauret (2007 : 29) : « [...] Le travail sur la prononciation est beaucoup plus difficile lorsque l'exposition à l'écrit a déjà produit une interlangue phonétiquement très marquée. Plus cette interlangue est pratiquée et acceptée comme 'du français', plus il est difficile et long de la remettre en cause. La 'souplesse' n'est pas facile à réacquérir. »

Autrement dit, les apprenants peuvent faire des progrès dans certains domaines (d'où une amélioration globale de 10%), mais certains aspects du système phonétique de leur interlangue sont fossilisés, tels que la prononciation des consonnes finales et en particulier la prononciation du /n/ en position finale comme dans *un, aucun*, etc.

En ce qui concerne la fossilisation des apprenants anglophones, Thomas (2008 : 595) remarque par ailleurs qu'après avoir passé une année en France, ils ont assez peu amélioré leur vocabulaire si ce n'est dans l'acquisition du vocabulaire familier. Ils continuent à faire des fautes d'accord et de genre (les exemples donnés sont *son poche... son jambe ... un queue*) et les anglicismes persistent (fossilisation lexicale). Ceci souligne le phénomène de fossilisation qu'une année en France ne suffit pas à enrayer. Or, si l'auteur remarque des erreurs fossilisées de vocabulaire et de grammaire, et que 75% de ces erreurs proviennent de l'interférence de la langue maternelle, il n'est pas surprenant de trouver des phénomènes similaires dans la prononciation des apprenants.

Lauret (2007 : 29) explique ce phénomène : « Les approches communicatives qui visent à mettre l'apprenant immédiatement au contact de la langue sous ses diverses formes, écrites en particulier, et qui incitent dès le début de l'apprentissage à la production à visée communicative, en minimisant de fait le rôle de la prononciation, tendent probablement à empêcher un développement optimal des possibilités phonétiques. [...] »

Il existe donc des catégories phonétiques et/ou syntaxiques qui sont particulièrement problématiques pour les apprenants. En fait, on observe un cumul de plusieurs facteurs, tout comme pour les francophones. Si une séquence contient une structure syntaxique relativement peu fréquente de type [adj. + nom], dans laquelle l'adjectif contient un /n/, alors la liaison a peu de chances d'être réalisée correctement par les anglophones, comme par exemple dans *aucun effort*, où la liaison a été réalisée à 27% au temps 1 et à 37.8% au temps 3, alors qu'elle est réalisée à 100% par les francophones majoritaires.

Bien que le /n/ soit fréquent dans la liaison en français, c'est un phonème qui est aussi très fréquent en position finale en anglais *an, pin, pen, bin, etc.*, mais il est prononcé différemment en français et en anglais. On peut constater trois différences importantes dans l'articulation du /n/ final entre le français et l'anglais. Tout d'abord, le point d'articulation est différent : le /n/ est une consonne dentale en français, tandis que c'est une consonne alvéolaire en anglais. De plus, le /n/ en anglais nasalise la voyelle qui précède dans la même syllabe comme dans *can*, alors que ce n'est pas le cas dans la prononciation française du mot le plus proche : *Cannes*. Enfin, en français, il y a une tendance à avoir une détente consonantique quand le /n/ est en position finale, ce qui a pour résultat qu'on l'entend clairement. Par contre, en anglais, les occlusives en fin de mot sont souvent prononcées sans détente consonantique (*unreleased stop*), et la nasale peut même se réaliser comme une nasalisation de la voyelle précédente, et une articulation alvéolaire très légère (une approximante).

Bref, une telle différence de prononciation pourrait expliquer que le /n/ en position finale prononcé différemment du français, rende l'enchaînement plus difficile. Cela pourrait expliquer que les apprenants prononcent /okæn efor/ avec le /n/ prononcé sans détente consonantique, ce qui aurait pour effet de bloquer la liaison (la liaison a été correctement réalisée dans *aucun effort* dans seulement 27% des cas).

Ainsi, la fréquence des structures ne garantit pas qu'elles soient acquises, si la structure phonétique présente dans cette séquence est différente en anglais. L'interférence de la langue maternelle peut ralentir et limiter l'apprentissage de la prononciation de la langue cible parce que les locuteurs sont, selon les termes de Troubetzkoy (1939), « sourds » à ses éléments phonétiques, à cause d'un filtre phonologique créé par la L1.

Dans notre étude, l'interférence de la langue maternelle semble encore plus forte chez les anglophones que chez les autres groupes linguistiques étudiés, bien que les langues asiatiques telles que le chinois soient, comme le précise Lauret (2007, 35) : « parmi les langues les plus différentes des caractéristiques de la prononciation du français [...] avec leurs caractéristiques prosodiques, syllabiques et segmentales. » Les anglophones bilingues et les personnes parlant une L2 à haut niveau ont de meilleurs résultats que les anglophones monolingues, ce qui suggère que ces locuteurs trilingues bénéficient de l'apport d'une autre langue que l'anglais.

Si les locuteurs parlant une autre langue que l'anglais peuvent bénéficier de l'apport de l'autre langue, le phénomène inverse apparaît également. Les locuteurs d'une L1 ayant l'anglais comme L2 et le français comme L3 subissent l'influence « négative » de l'anglais sur la prononciation du français. Les liaisons et les enchaînements sont déjà difficiles pour tous les apprenants L2. Özçelik (2008) remarque que les étudiants turcs du français ont des problèmes aux niveaux segmentaux et suprasegmentaux. Ils ont les problèmes communs à tous les apprenants du français langue seconde ou étrangère, mais ils ont en plus des

problèmes d'interférence de leur langue maternelle et de leur langue étrangère première : l'anglais. Cela suggère donc que l'apprentissage de la prononciation du français pourrait être plus difficile pour les anglophones, mais pourrait l'être aussi pour les apprenants ayant l'anglais comme L2. Par contre, si tous parlent l'anglais (comme dans notre étude), l'apport d'une troisième langue, compenserait, en partie, les difficultés causées par l'interférence de l'anglais.

D'un autre côté, une autre explication pourrait être que faute d'enseignement de la prononciation, et étant donné un input trop restreint pour apprendre le système phonétique « naturellement », les apprenants trilingues ont recours aux systèmes qu'ils connaissent mieux, le système phonétique de leur L1 et de leur L2. Le transfert de ces langues ne serait alors non pas la cause, mais plutôt une conséquence d'un autre problème : le manque d'input et l'absence d'enseignement de la prononciation.

En effet, Özçelik observe que l'enseignement de la prononciation est clairement insuffisant et qu'il faut trouver des façons de remédier à ces lacunes. Selon l'auteur (p. 206) les apprenants ont besoin de pratiquer la coordination des symboles écrits [graphèmes] et des sons [phonèmes], la reconnaissance des mots au plan sémantique, la segmentation du discours, l'intonation, l'accent, le rythme, l'enchaînement, la liaison, la vitesse de la lecture, et de palier au manque de connaissances sur la ponctuation, de pratiquer la prononciation du /R/ « uvulaire » et de se détacher de l'influence de l'anglais.

Stridfeldt (2002) souligne que les liaisons et les enchaînements sont un problème important pour les étudiants suédophones de l'étude (et tous les autres), parce que les frontières de mots écrits ne correspondent pas aux frontières de syllabes. De plus, alors qu'en anglais (ou d'autres langues) les accents de mots permettent de fournir d'autres indices sur les mots, ce n'est pas le cas en français où l'accent rythmique se trouve à la fin des groupes de mots, au détriment de mots isolés. Les différences entre graphie et phonie participent aux

difficultés des apprenants qui sont négativement influencés par l'écrit et les frontières de mots écrits qui ne correspondent pas aux frontières de mots à l'oral.

Berri (2007) déplore le manque d'enseignement de la prononciation du français dans le sud du Brésil au niveau universitaire, et les raisons et conséquences de cette lacune. Il prétend qu'une mauvaise prononciation dissuade les étudiants de parler (surtout si elle est stigmatisée), ce qui cause des problèmes dans tous les autres aspects langagiers, tels que la grammaire. C'est la raison pour laquelle l'apprentissage de la prononciation est une condition *sine qua non* pour l'acquisition des autres composantes linguistiques.

Syrovatskaja (2008), professeure de français en Lituanie, regrette le manque de références phonétiques dans les manuels de langue, car selon elle, l'important dans la communication, c'est d'abord le rythme et la prosodie. L'auteure pense que les erreurs de prononciation empêchent la communication avec des locuteurs natifs parce qu'elles attirent l'attention des auditeurs sur des particularités incorrectes de la prononciation, sans parler de la compréhension erronée de l'énoncé. Dans certains cas de communication, ce qu'on dit n'a pas une grande importance, car l'important c'est le ton et le timbre de la voix.

Syrovatskaja (2008 : 3) préconise un travail articulaire en début d'acquisition afin que l'apprenant s'approprie une prononciation figée et irréversible, affine son oreille, développe sa souplesse perceptivo-motrice et la fluidité de la parole, ce qui lui permettra de s'orienter avec aisance dans le continuum sonore de la langue étrangère et lui permettra d'apprendre la langue dans toute sa totalité émotionnelle, affective et culturelle.

Dansereau (1995) considère que la phonétique devrait être enseignée en début d'apprentissage, de façon systématique et graduelle avec de nombreuses répétitions pour que les apprenants prennent de bonnes habitudes et n'oublient pas ce qu'ils ont appris. Elle conseille aussi de traiter de la conversation dans différentes tâches telles que les exercices et règles de grammaire (masculin/pluriel, féminin/masculin) et de parler aussi de la prosodie, en



commençant avec l'accent tonique avec des sons des mots similaires (*papa, Canada*), puis la syllabation (les liaisons et enchaînements) et l'intonation.

Selon l'auteure, après avoir au préalable été formés en début d'acquisition, les étudiants devraient lire en classe, être corrigés et avoir l'occasion de pratiquer les règles apprises avec des exercices oraux et des exposés. Si les étudiants font attention à la prononciation, et que le professeur fait souvent remarquer des éléments phonétiques aux étudiants, ceux-ci seront habitués à la discrimination auditive et auront ainsi de meilleures compétences d'écoute et de production orale.

La quasi-totalité des chercheurs s'accordent à dire que l'enseignement de la prononciation est une nécessité, et les études portant sur les effets de l'enseignement de la prononciation en indiquent les effets positifs et parfois même spectaculaires. Dans notre étude, la petite leçon de phonétique dans les quatre groupes suggère que la fréquence et la répétition jouent un rôle non seulement indispensable dans l'acquisition d'items et de structures phonétiques, mais que cette répétition serait même plus importante que les autres éléments d'une leçon de phonétique, à savoir explications et corrections.

En fait, toutes les situations d'enseignement offrant aux apprenants un input riche et varié, une sensibilisation aux éléments suprasegmentaux et aux éléments différents de la L1, une écoute intensive et beaucoup de répétitions permettent toutes aux étudiants de faire des progrès. Que ce soit notre demi-heure de cours avec en particulier l'écoute et la répétition, où les nombreuses autres études portant sur l'enseignement de la phonétique.

Catford et Pisoni (1970) concluent que le groupe ayant reçu des informations explicites sur le système articulatoire de la langue cible a obtenu de meilleurs résultats que le groupe ayant reçu seulement des renseignements sur comment prononcer certains mots ou phrases.

Leather (1990) a prouvé que les imitateurs naïfs des tons du mandarin ont eu de meilleurs résultats lorsqu'ils ont suivi au préalable un entraînement sur la perception des tons, que lorsqu'ils étaient d'abord entraînés à la production et ensuite à la perception.

Elliott (1994), dans son étude portant sur des apprenants anglophones apprenant l'espagnol, a montré que l'enseignement de la prononciation permet de faire des progrès avec les éléments segmentaux et suprasegmentaux dans la répétition comme dans la parole spontanée.

Ashby, Maidment et Abberton (1996) ont instauré un système d'apprentissage basé sur l'écoute et la discrimination avec des Questions à Choix Multiples, afin d'habituer les oreilles des apprenants à discriminer les éléments les plus simples comme les plus subtils. Cette approche a permis « d'affiner » petit à petit les oreilles des apprenants jusqu'à ce qu'ils soient capables de percevoir des différences très subtiles et les résultats ont prouvé la supériorité de la perception sur la production.

Markham et Nagano-Madsen (1996) soutiennent que, si les sujets ont été entraînés à imiter les sons d'une langue étrangère, ils ont généralement de bons résultats en ce qui concerne leur capacité à imiter la langue cible et peuvent produire des versions authentiques des stimuli.

Akahane-Yamada, Tohkura, Bradlow et Pisoni, (1996) soulignent qu'un entraînement en perception produit des effets positifs à la fois sur la perception (la capacité de discrimination) et sur la production des apprenants japonais avec les paires minimales /r/ et/l/ qui sont connues pour leur causer de grandes difficultés, alors qu'il n'y a eu aucune différence pour le groupe témoin. Les tests effectués trois mois et six mois plus tard, indiquent que l'effet est encore plus bénéfique à long terme que juste après l'entraînement.

Abe (2006) a étudié les effets de l'instruction formelle de la prononciation sur la perception et la production d'étudiants japonais apprenant l'anglais. Les deux groupes ayant

reçu de l'instruction phonétique explicite ont eu de meilleurs résultats que le groupe témoin.

Knoerr et Weinberg (2005) ont comparé deux approches d'enseignement de la prononciation dans le cadre d'un cours universitaire de français langue seconde de niveau élémentaire faible : avec une approche traditionnelle avec des cassettes audio, et avec une approche multimédia à l'aide d'un cédérom. Les résultats suggèrent que les deux groupes ayant reçu une formation en phonétique ont fait des progrès et plus précisément :

L'analyse qualitative a mis en évidence le fait que le groupe qui a fait le plus de progrès était celui qui avait reçu l'approche combinée ordinateur et enseignant. [...] L'étude a révélé que la performance des apprenants recevant un enseignement explicite de la phonétique en FLS était supérieure à celle des apprenants n'en recevant aucun. Elle a également indiqué qu'il n'y avait pas de différence significative au niveau de l'apprentissage entre le groupe traditionnel et le groupe multimédia. Par contre, le niveau de satisfaction était plus élevé chez les étudiants du groupe multimédia. (Knoerr et Weinberg, 2005 : 385-386)

Bastin (2007) a créé une méthode (la méthode *Balingua*), basée sur l'oral et sur la façon dont les enfants apprennent à parler. L'apprenant écoute de nombreuses fois des phrases, surtout des questions et des réponses sur l'ordinateur, puis il participe à une pseudo conversation : on lui pose des questions et il y répond. Le vocabulaire s'apprend à l'aide de symboles et l'écrit n'est vu qu'une fois l'oral appris. Apparemment la méthode a fait ses preuves et de nombreux apprenants venant du monde entier ont appris des langues très différentes. Ce n'est pas un cours de langue complet dans la mesure où il y a relativement peu d'heures et d'explications, mais cela permet de prendre un bon point de départ avec une bonne prononciation.

Yates (2003) explique les avantages de l'imitation de cassettes vidéo pour améliorer la prononciation des apprenants, et améliorer les capacités d'imitation en particulier avec les éléments suprasegmentaux

Knoerr (2000) a étudié les effets de l'utilisation du logiciel *Videovoice* créé et utilisé à l'origine pour les sourds de naissance et en a conclu que le logiciel est très efficace pour l'enseignement de l'intonation. L'intonation était au départ un grand problème pour les 27

étudiants. Cette étude montre que les progrès sont nets, et que c'est justement dans les aspects les plus difficiles qu'il y a le plus de progrès. Les étudiants qui ont le plus pratiqué ont fait le plus de progrès.

En outre, si les apprenants peuvent faire des progrès par la suite, c'est surtout en début d'acquisition que les apprenants ont le plus besoin de répétitions et de pratique, parce que c'est à ce moment-là que se forment les habitudes en prononciation. Comme le précise Lauret (2007) :

C'est en effet particulièrement en début d'apprentissage, que la prononciation quasi conforme au modèle standard « d'un répertoire très limité d'expressions et de mots mémorisés » semble possible, en particulier via la mémorisation / reproduction de modèles qui fait plus appel à l'auditif qu'au linguistique. Néanmoins, cette performance de mémorisation et d'imitation peut être considérablement réduite par l'exposition à l'orthographe, la faible disponibilité de modèles, le manque de pratique, par exemple. (Lauret, 2007 : 154)

Boubnova (2000) a étudié les effets d'une instruction formelle et d'un entraînement phonétique intensif en début d'acquisition. Des apprenants débutants ont reçu une formation de quatre semaines à raison de quatorze heures par semaine. L'enseignement est basé sur une étude contrastive, c'est-à-dire sur ce qui est le plus différent entre les deux langues (français et russe). Comparé à un groupe témoin, et surtout au test fait avant la formation, les résultats sont spectaculaires. Les apprenants arrivent à prononcer les phrases de façon native.

Si l'apprentissage de la prononciation se fait au début, et qu'en plus de cela, il y a une formation intensive sur des aspects isolés (comme ici l'accentuation de syllabe, la prononciation de la syllabe), avec de nombreux exercices qui réunissent des aspects visuels et auditifs, les progrès sont énormes. Les apprenants prononcent alors de façon spontanée et native, leur prononciation native devient un réflexe.

Enfin, les résultats les plus spectaculaires sont, à notre avis, ceux des études de Neufeld (1979, 1980). 20 sujets ont reçu 18 heures de cours de prononciation, dont la moitié en chinois et l'autre en japonais. Durant la première partie du programme, ils ne faisaient

qu'écouter, discriminer des sons et des intonations allant du plus simple au plus subtil. Pendant cette partie de la formation, ils devaient seulement écouter.

Dans la deuxième partie de la formation, ils ont commencé à chuchoter puis ils ont imité des sons, des mots et des phrases allant d'une syllabe jusqu'à 16 syllabes. Enfin, ils ont fait un test d'imitation qui consistait à répéter une dizaine de phrases jamais entendues auparavant. Leur enregistrement a été analysé par trois juges natifs pour le chinois et trois juges natifs pour le japonais. 11 des 20 sujets ont été identifiés comme étant de langue maternelle japonaise et 9 comme étant de langue maternelle chinoise. Les 20 sujets n'avaient eu comme seule formation qu'un entraînement à la discrimination phonétique et à la production de 9 heures dans chaque langue. La moitié des apprenants ont une prononciation native dans des langues réputées très difficiles, en ayant eu seulement 9 heures de cours.

Malgré les résultats révélateurs de nombreuses études, certains chercheurs ne pensent pas que les cours de phonétique soient une condition primordiale dans l'acquisition de la prononciation. D'après leur étude portant sur 61 apprenants non-natifs de l'anglais, Purcell et Suter (1980) concluent que les variables qui prédisent le plus une bonne prononciation (*accuracy*) en L2 sont la L1 et la proximité de la L1 par rapport à la L2, une bonne capacité à imiter (*good mimics*), le fait d'avoir vécu dans un pays parlant la langue cible pendant plusieurs années, d'avoir habité avec des locuteurs natifs et d'avoir le désir d'améliorer sa prononciation. Purcell et Suter (1980) précisent que ces variables ne dépendent pas de l'enseignement de la L2 en salle de classe.

Nous supposons qu'ils entendent par cela des cours traditionnels de langue (dans lesquels la prononciation n'est pas enseignée), car la plupart des études, y compris la nôtre, suggèrent autre chose. Même si nous sommes d'accord que la L1 joue un rôle indéniable (ici défavorable pour les anglophones apprenant le français), nous soutenons que la L1 n'est pas le facteur le plus important. Par exemple, les Chinois de notre étude parlent une langue (dont

certaines sont bilingues) la plus éloignée du français parmi tous les locuteurs et ils ont pourtant eu des productions de liaisons parmi les plus élevées. Par contre, ce qui est intimement lié à une L1, c'est le système d'éducation, les approches pédagogiques utilisées dans le système scolaire et universitaire, ainsi que l'input reçu et les conditions d'apprentissage.

Ainsi, des locuteurs d'une même L1, peuvent avoir des résultats très différents si les conditions d'apprentissage sont différentes. Par exemple les 20 sujets anglophones de Neufeld (1979, 1980) apprenant le chinois et le japonais avec les approches ci-mentionnées ont obtenu de meilleurs résultats en prononciation que ceux ayant appris ces mêmes langues dans des cours de langue « traditionnels » ou en ayant même séjourné dans les pays où l'on parle les langues cibles (aucune autre étude fait état de 50% d'apprenants adultes parvenant à prononcer le chinois et le japonais sans accent après seulement 9 heures de cours).

Paradoxalement, Purcell et Suter (1980) considèrent qu'une bonne capacité à imiter la L2 est l'un des facteurs les plus importants pour prédire le succès des apprenants, alors que les cours de langue jouent un rôle très faible, voire inexistant, dans le succès des apprenants à prononcer de façon native. Or, les méthodes d'enseignement dans lesquelles sont présentes en début d'acquisition l'écoute intensive et la répétition ont justement pour but de mener les apprenants à une excellente imitation, et ce sont précisément ces éléments qui peuvent le mieux prédire le succès des apprenants. L'écoute et la répétition influencent fortement la capacité à imiter. Mais plus que répéter, les apprenants doivent apprendre à imiter. Imiter est plutôt comme parodier la langue cible en distinguant surtout les éléments saillants, différents de la langue maternelle et en les exagérant (Lauret, 2007 : 88). Les apprenants L2 bénéficient le plus de la répétition et de l'imitation, parce que seules ces activités permettent l'acquisition définitive de structures qui deviennent alors des réflexes, ce que mentionne Lauret (2007) :

Acquérir une nouvelle prononciation nécessite toujours de l'entraînement, même pour les personnes « douées ». [...] En prononciation, comme pour un sport ou une performance physique, motivation et entraînement sont les maîtres mots. Il y a finalement peu de recours « intellectuels » pour aider à obtenir cette performance orale. L'esprit, l'oreille et la voix doivent s'ouvrir positivement à la nouvelle langue et s'entraîner. (Lauret, 2007 : 25)

Ainsi, la partie de notre étude portant sur le rôle de la répétition dans des cours de prononciation s'accorde avec les études dans le domaine. Un input riche et varié permet l'acquisition d'une L1 et l'acquisition d'une L2 dans les écoles francophones. Par contre, le manque d'input en début d'acquisition handicape fortement les apprenants et limite les progrès qu'ils pourront faire par la suite.

Contrairement à l'acquisition du lexique, de la morphologie ou de la syntaxe d'une L2, l'acquisition de la phonétique ne se développe pas de façon linéaire. Elle s'opère très rapidement en début d'acquisition pour la grande majorité du système phonétique, sous forme de réflexes, réflexes majoritairement conditionnés par la L1, que seules une écoute et une pratique intensives permettent d'acquérir correctement (Lauret, 2007 : 155).

## **CONCLUSION**



Le but de cette thèse était d'analyser certains facteurs qui contribuent à la production de liaisons obligatoires des anglophones et accessoirement d'analyser les liaisons facultatives et interdites dans le groupe nominal. Le Modèle Basé sur L'Usage permet de rendre compte de nombreux phénomènes qui semblent au premier abord n'avoir aucun lien entre eux (les erreurs des enfants et des adultes, l'acquisition des liaisons chez les enfants L1, la production des adultes L1 et la production des apprenants L2), mais qui en réalité traitent tous du même phénomène sous des angles différents : le rôle de la fréquence dans l'acquisition et la production de liaisons.

Nous avons analysé divers facteurs pour comprendre pourquoi certaines liaisons sont obligatoires, certaines sont facultatives et d'autres interdites dans le même contexte phonétique et syntaxique (déterminant + nom, adjectif + nom ou expression figée). À l'intérieur d'une même catégorie, les séquences des sous-catégories ne sont pas traitées de la même façon. Parmi les liaisons obligatoires, facultatives et interdites, certaines sont produites plus que d'autres.

Notre étude constate, en accord avec, entre autres, Laks (1977), Encrevé (1988) et Durand et Lyche (2008), que les francophones ne font pas de liaisons systématiques avec les adjectifs masculins antéposés au nom (94.8%), à l'exception de quelques adjectifs les plus fréquents. Les liaisons avec les adjectifs qualificatifs sont donc des liaisons facultatives très fréquentes, ainsi que les liaisons avec les adjectifs ordinaux + nom, réalisées à 70.8% par les francophones majoritaires, contrairement aux descriptions de l'Académie Française. Il en est de même avec les expressions figées, dans lesquelles les liaisons sont réalisées à 87.5% par les francophones majoritaires et les liaisons interdites qui sont réalisées (prononcées) dans 13.7% des cas par les francophones majoritaires.

Il apparaît donc que la classification *liaisons obligatoires, facultatives et interdites* est problématique, parce que de telles séparations sont artificielles et ne représentent pas la réalité

linguistique. Mais si cette classification est imprécise, comment rendre compte du fait qu'à l'intérieur de certaines catégories, les liaisons avec certaines séquences ou avec certains mots sont toujours réalisées alors que d'autres non ? Il paraît plus juste de parler d'un continuum des liaisons les plus fréquemment réalisées, aux liaisons les plus rarement réalisées. Ou bien de changer la terminologie et plutôt qu'utiliser le terme « obligatoires » utiliser un autre terme tel que « très fréquentes ».

Les résultats de cette étude confirment nos hypothèses de départ. En ce qui concerne les productions des francophones, plusieurs facteurs rentrent en jeu pour « la validation » (*entrenchment*) de structures et les productions de liaisons : la fréquence lexicale (*token frequency*), la fréquence des structures syntaxiques (*type frequency*) (Bybee, 2001), la productivité et la complexité des structures syntaxiques, dont le nombre de mots qu'englobe une structure (d'après le « Modèle de la Concurrence » de MacWhinney, 1978, 1982, 1986, 1987 et de Bates et MacWhinney, 1979, 1982), et la fréquence et la productivité des consonnes de liaisons (/n/, /t/, /z/, /R/) (Mallet, 2008).

Dans leur apprentissage des liaisons, les enfants francophones passent par trois étapes (Dugua, 2005) qui correspondent à la construction d'une banque de données donnant progressivement naissance à des catégories abstraites permettant aux enfants d'utiliser les données enregistrées pour les généraliser et les appliquer à des données nouvelles jamais entendues auparavant. Cela crée d'ailleurs quelques erreurs qui vont progressivement s'estomper, mais quelques-unes vont persister dans le lexique adulte, en particulier chez les adultes les moins éduqués en situation d'hypercorrection.

Les liaisons sont acquises dans un ordre prévisible. Les enfants apprennent les liaisons obligatoires avant les liaisons facultatives (Chevrot, 2002 ; Chevrot et Fayol, 2000 ; Chabanal, 2003), parce que les liaisons obligatoires sont plus fréquentes dans leur input que les liaisons facultatives. Le rôle de l'input est visible dans la production de liaisons d'enfants de tous les

âges étudiés, ce qu'on voit d'après la classe socioéconomique des parents (Chabanal et Embarki, 2002 ; Chabanal, 2007).

Les erreurs que font les enfants francophones sont également prévisibles car elles dépendent de la fréquence des séquences de leur input (Dugua, 2006). On retrouve un processus similaire avec les erreurs de liaisons, de morphologie et de syntaxe (Pallaud et Savelli, 2001 : 130-131). Les enfants utilisent un processus d'analogie qui consiste à uniformiser, simplifier et surgénéraliser les règles connues des items souvent entendus aux items peu ou jamais entendus. Ils appliquent les règles des classes les plus productives aux classes les moins productives (par exemple deux-z-héros). Les enfants rajoutent une consonne d'épenthèse qui correspond aux séquences les plus proches les plus souvent entendues et font encore ce genre d'erreurs jusqu'à l'âge de 11-12 ans (Pallaud et Savelli, 2001).

Le type d'erreurs que font les enfants qui ont atteint la dernière étape de développement (6-7 ans) est d'ailleurs similaire aux erreurs que font les adultes et se font approximativement dans les mêmes contextes : cuirs et velours avec formes verbales, qui sont les structures les plus rares et les moins productives parmi les liaisons facultatives. Les erreurs des adultes sont tout aussi prévisibles. Ils ne font jamais d'erreurs dans les contextes les plus fréquents de type [dét. + nom], [pron. + pron.], [pron. + verbe ni même avec les liaisons facultatives fréquentes de type [adv. de négation (pas) + part. passé], [être + adj.] ou [être + part. passé], car ces structures sont acquises en premier, et sont parfaitement maîtrisées.

Dans un même contexte syntaxique et phonétique, certains adjectifs antéposés sont dénasalisés (*Le Moyen-Âge, bon ami*), d'autres ne le sont pas (*mon ami*) et avec certains adjectifs il n'y a pas de liaison du tout (*hautain, malin enfant*). Les différentes prononciations correspondent à différents niveaux de fréquence lexicale, fréquence du mot1, du mot2 et fréquence de cooccurrence des mots. Par exemple, non seulement les adjectifs *hautain* et *malin* sont rares, mais leur taux de cooccurrence est encore plus rare, il n'y a donc pas de

liaison dans *malin enfant*, car on dirait plutôt *un enfant malin*. Lorsque certaines personnes n'enchaînent pas les liaisons pour des raisons stylistiques (Encrevé, 1988), ce sont là aussi dans des contextes bien précis, dans des cas de liaisons facultatives, cas qui suscitent des erreurs chez les enfants comme chez les adultes.

La fréquence des séquences détermine aussi l'évolution des langues (Diessel, 2007 ; Ellis, 2002 ; Bybee, 2001a, 2001b). Comme nous l'avons vu dans les chapitres précédents, la majorité des mots des liaisons obligatoires appartiennent à la catégorie K1 et sont donc des mots fréquents. De plus, la majorité des structures syntaxiques des liaisons obligatoires sont des structures fréquentes et productives. En ce qui concerne les liaisons interdites de notre étude, 10 mots sur 13 mots sont des mots « rares » notamment les mots commençant par une semi-consonne [j] et par un h muet. Par contre, les liaisons avec les mots fréquents commençant par une semi-consonne (*yeux*) et un h muet (*hôpital, histoire*) appartenant à la catégorie K1 sont généralement classées comme liaisons obligatoires, alors qu'avec les mots plus rares les liaisons sont interdites (les mots commençant par un h aspiré et un y, sont des mots d'origine étrangère, généralement des mots rares).

Sur 51 séquences de liaisons obligatoires dans le texte de notre étude, cinq appartiennent à la catégorie K2 : *mes ouvriers, moyen-âge, certain âge, grand immeuble* et *folles aventures* (les deux mots appartiennent à la catégorie K2) ; deux appartiennent à la catégorie K3 : *fait accompli* et *ancien édifice* ; et cinq à la catégorie hors liste : *belles acclamations, vieux habits, vieilles éponges, léger ennui* (*léger* appartient à la catégorie K2 et *ennui* à la catégorie hors liste) et *accent aigu* (*accent* appartient à la catégorie K3 et *aigu* à la catégorie hors liste).

*Accent aigu* et *léger ennui*, qui sont les deux séquences les moins fréquentes des liaisons obligatoires, en plus d'appartenir à des catégories moins productives que les autres, sont précisément celles dont la liaison est la moins réalisée par les francophones majoritaires.

La liaison avec *accent aigu* est réalisée à 58.3% et avec *léger ennui* à 41.7%.

L'analyse de la fréquence des mots1 et des mots2 avec *GoldVarb* à partir de *Vocabprofil* et de *Lexique* a montré que les productions de liaisons des francophones majoritaires et minoritaires et des apprenants du français L2 sont proportionnelles à la fréquence des mots1, des mots2 et à la fréquence de cooccurrence des mots1 et mots2.

En plus de la fréquence des items, la fréquence et la productivité des structures jouent un rôle tout aussi important. Le nombre de mots appartenant à chaque classe (ou catégorie), participe à la productivité d'une structure. Plus une structure contient d'items, plus elle est productive. Le cerveau humain ayant enregistré un grand nombre d'items valide et renforce une structure souvent entendue et utilisée.

Un autre facteur qui participe à la productivité d'une structure est la complexité d'une règle : le nombre de séquences qui appartiennent à cette catégorie, par rapport au nombre de séquences qui n'y appartiennent pas. Ainsi, si une structure contient peu de mots, mais que dans cette même structure syntaxique ou phonétique, il n'y a pas de cas contraires, alors la règle est relativement simple à traiter. Par exemple, pour les francophones, le cas de *neuf* où le f se prononce /v/ est relativement simple à traiter, puisqu'il n'y a que deux cas : [*neuf + heures*], et [*neuf + ans*]. Dans les autres cas, le f se prononce /f/ (*neuf enfants*), et c'est alors un enchaînement normal. C'est la raison pour laquelle la liaison dans *neuf heures* a été réalisée correctement à 100% par tous les francophones.

Par contre, lorsque des items sont moins fréquents et qu'en plus, ces items vont à l'encontre des règles générales, alors la structure est moins productive et les francophones font moins de liaisons. C'est le cas notamment des expressions figées où *fait accompli* a été réalisé à 83.3% et *accent aigu* à 58.3%. Si une structure est peu productive, mais que l'item de cette structure est fréquent, la fréquence permet de compenser le manque de productivité, comme c'est le cas avec *États-Unis* (dont la liaison a été réalisée à 100% par tous les

francophones), séquence dans laquelle la liaison est réalisée dans une plus grande proportion que *accent aigu* (dont la liaison a été réalisée à 58.3% par tous les francophones).

Toutefois, même si *États-Unis* est une séquence fréquente, sa fréquence ne permet pas de compenser le manque de productivité de la structure. C'est pourquoi un ou deux items très fréquents dans une classe où le reste des items sont peu fréquents a pour effet, qu'au total la structure est peu productive. En effet, il y a plus de cas où on ne fait pas la liaison dans le contexte [nom + adj.], que de cas où l'on fait la liaison. De même, on fait plus de liaisons dans le contexte [nom + adj. pluriel], que dans le contexte [nom + adj. singulier]. Par contre, ce qui explique la réalisation à 100% de *moyen-âge*, n'est pas tant sa fréquence, mais c'est plutôt sa relative productivité syntaxique. Il est plus fréquent de faire une liaison dans le contexte [adj. + nom] (*ancien élève*), que dans le contexte [nom + adj. singulier], même si le contexte [adj. + nom], n'est pas une classe très productive par rapport à d'autres car elle comprend peu d'items.

Enfin, les consonnes de liaison jouent également un rôle, puisqu'elles constituent des classes phonétiques. Comme avec les catégories syntaxiques, les francophones les traitent selon leur productivité, ce qui comprend le nombre d'items et de structures de cette catégorie. Par exemple, le nombre de structures permettant la liaison avec un /R/ sont rares : [adj. singulier + nom] (*léger ennui*), [infinitif + prép.] ou [infinitif + adverbe] (*arriver\_à l'heure*, *aller\_encore*). Les séquences [adj. singulier se terminant par r + nom] sont encore plus rares: *premier*, *dernier*, *léger*, peut-être quelques autres de plus.

La séquence [adj. + nom] est déjà relativement rare puisque, comme nous l'avons vu dans les chapitres précédents, la plupart des adjectifs se placent après le nom, et s'ils se placent avant, ils se terminent très rarement par un r. Et si c'est le cas, les séquences qui réunissent les trois paramètres : adjectif antéposé, adjectif se terminant par r et nom commençant par une voyelle, sont encore plus rares. La structure syntaxique est relativement

rare et la structure phonétique est particulièrement rare. Cela a pour conséquence que les liaisons avec /R/ sont les plus rarement réalisées, par rapport aux autres consonnes de liaisons : 77.8% réalisées par les francophones majoritaires et 25% par les francophones minoritaires.

Ainsi les locuteurs francophones traitent simultanément des informations syntaxiques, phonétiques, lexicales et stylistiques pour calculer intuitivement des statistiques qui se basent sur une banque de données très riche. Ils établissent ainsi une hiérarchie des liaisons qui doivent être réalisées en priorité.

Bybee (1995), Bybee et Thompson (2000) et Bybee et Hopper (2001) ont montré que la productivité des modèles phonologiques, morphologiques et syntaxiques dépend surtout de la fréquence des constructions plutôt que de la fréquence des mots. En effet, l'apprentissage langagier commence par l'acquisition d'items individuels (Tomasello, 2000a, 2000b; Ellis, 2002) et une fois un nombre assez important d'items acquis, la fréquence des constructions est plus importante que la fréquence des items.

Nous avons constaté cela dans notre étude. Par exemple, la séquence *un accent aigu* demande deux liaisons obligatoires : l'une après le déterminant *un*, et l'autre dans l'expression figée *accent aigu*. Il s'agit de la même séquence, donc de la même fréquence lexicale (du mot<sup>2</sup>), fréquence rare. La liaison avec le déterminant a été réalisée à 100%, alors qu'elle n'a été réalisée qu'à 58.3% dans *accent aigu*. Cela souligne l'importance de la fréquence syntaxique.

Toutefois si les structures syntaxiques sont rares, c'est alors la fréquence lexicale qui détermine la production de liaison. C'est pourquoi on ne peut pas vraiment séparer la fréquence lexicale de la fréquence syntaxique, parce que l'une dépend de l'autre, et le tout dépend de la productivité des structures syntaxiques et phonétiques. C'est la raison pour

laquelle on ne peut comprendre et prédire les productions de liaisons que si l'on considère tous ces facteurs liés les uns aux autres.

Nous nous sommes concentrés principalement sur les liaisons obligatoires dans le groupe nominal, en particulier dans le contexte [dét. + nom], et accessoirement [adj. + nom] et avec les expressions figées dans le groupe nominal. Il serait utile d'analyser les autres contextes de liaisons obligatoires : les liaisons avec un pronom (pronom personnel + verbe, pronom + pronom, etc.), les liaisons avec une préposition monosyllabique, une conjonction monosyllabique (quand) et un adverbe monosyllabique (*très*, *bien*) et d'autres expressions figées.

Il serait aussi utile d'analyser les liaisons facultatives et les liaisons interdites plus en détail. Pour cela, il faudrait étudier des contextes plus variés et surtout avoir un plus grand nombre de locuteurs.

En ce qui concerne les locuteurs du français L2, il faudrait aussi avoir un nombre de locuteurs plus élevé, ainsi qu'un plus grand nombre de locuteurs dans chaque sous-catégorie. Par exemple, si l'on veut étudier l'influence de la L1, ou les effets du bilinguisme, il faudrait avoir des groupes et sous-groupes linguistiques plus grands.

Notre mini-formation durait 30 mn et notre post-test a eu lieu une semaine après la leçon, à cause des impératifs du cours. Si l'on veut étudier les effets de l'enseignement sur des points phonétiques tels que les liaisons, il serait utile d'avoir une formation plus longue (qu'une demi-heure), de préférence une année ou deux années d'enseignement, et d'avoir un post-test beaucoup plus tard que le nôtre, plusieurs mois plus tard, voire une année plus tard.

Nous avons constaté les effets de la fréquence chez les apprenants L2. Plus les apprenants ont eu un input riche, plus ils réalisent de liaisons (ceux issus du système francophone ou du français intensif). Certes, nous n'avons qu'un seul sujet ayant étudié quelques années en école francophone ; mais le simple fait que la vaste majorité des étudiants



issus de ce système aillent en deuxième et même en troisième année (d'après le test de placement) de spécialisation en français, tout comme les francophones majoritaires soutient cette hypothèse.

Si l'input des apprenants est inférieur à celui des francophones, leur production de liaisons l'est aussi. Parmi le groupe de francophones, on remarque également le rôle crucial de l'input. Ainsi, les francophones majoritaires font plus de liaisons obligatoires (95.6%) que les francophones minoritaires (85.5%), qui en font eux-mêmes beaucoup plus que les anglophones (60.7% au temps 1).

Une analyse plus approfondie a permis d'expliquer les principales différences entre les groupes francophones et anglophones. Alors que les francophones se basent sur des facteurs syntaxiques, phonétiques et lexicaux pour calculer des « statistiques », les apprenants n'ont pas les données suffisantes qui leur permettent d'analyser les informations de manière aussi performante que les francophones. Les apprenants L2 semblent se baser principalement sur la fréquence lexicale, et cette fréquence lexicale est limitée, pour la plupart, à l'input reçu en salle de classe.

Comme les apprenants n'ont pas reçu la fréquence nécessaire dont ils avaient besoin en début d'acquisition, ils ont formé une interlangue dont de nombreuses structures et séquences sont fossilisées. Ils ont quand même pu apprendre certains items et certaines séquences (d'où les 60% de productions de liaisons obligatoires), mais ces connaissances ne leur permettent pas d'appliquer des règles abstraites à des items peu connus ou inconnus, de façon systématique, comme le font les francophones.

Les liaisons ne représentent qu'une petite partie de la prononciation, mais elles n'en sont pas moins représentatives de la prononciation des apprenants. Nous n'avons pas abordé ici les autres aspects phonétiques, mais en écoutant les enregistrements, nous avons observé que la liaison n'était qu'un problème parmi d'autres (intonation, accentuation de syllabe,

égalité syllabique, etc.). La prononciation des consonnes finales et la prononciation d'éléments segmentaux en général est l'un des problèmes de prononciation parmi d'autres.

Par exemple, il y a eu 98 cas (sur 1887) de consonnes de liaison mal prononcées et 28 cas de phonèmes ou syllabes du mot d'avant mal prononcés. Cela ne comprend pas le /R/, le /l/ ou le /n/ prononcés comme en anglais ou les voyelles diphtonguées. Nous n'avons compté comme erreur que si un phonème est prononcé comme un autre phonème existant en français, par exemple si un /z/ est prononcé /s/.

Il y a aussi eu 4 cas de phonèmes ou syllabes du mot d'après mal prononcés. Cela constitue un total de 130 phonèmes mal prononcés sur 1887, soit 6.89% des erreurs, en plus des 167 consonnes finales prononcées non-enchaînées soit 8.85% des cas. Il s'avère donc que 297 phonèmes ont été mal prononcés d'une façon ou d'une autre, sur 1887 séquences, soit un total de 15.74%. À cela s'ajoutent les pauses, hésitations, reprises avec ou sans liaisons, les cas de fausses liaisons et autres facteurs empêchant les liaisons.

De nombreux étudiants ont prononcé *ouvriers* /uvriɛr/, *assistants* /asistât/, *effort* /efɔrt/, *vieilles* /vjɛl/ ou /vjø/, *une* /œn/, *oies* /oj/. Si les étudiants prononcent mal les consonnes finales, et surtout prononcent des consonnes finales muettes, comme par exemple *effort* /efɔrt/, ou *vieilles* /vjɛl/ ou /vjø/, il est évident qu'ils ne pourront pas faire correctement les liaisons correspondantes.

Bien que dans les manuels de phonétique les liaisons soient traitées comme un phénomène séparé des éléments segmentaux et suprasegmentaux, nous considérons que les liaisons sont en fait directement reliées aux éléments segmentaux et suprasegmentaux. D'ailleurs, tous les éléments sont liés les uns aux autres : l'intonation, l'accentuation de syllabe, le découpage en groupes syntaxiques et en groupes syllabiques et la prononciation d'éléments segmentaux. C'est la raison pour laquelle, si les étudiants n'ont pas appris de

façon formelle les principales règles phonétiques, ils vont avoir des difficultés dans tous ces domaines qui eux-mêmes en engendrent d'autres.

Même si les étudiants connaissaient les règles sur les liaisons, qu'ils ont entendu de nombreuses fois des séquences en contenant (dans un cours de langue par exemple), mais qu'ils ne connaissent pas les règles d'accentuation de syllabe et de prononciation des consonnes finales, ils risquent de faire des erreurs dans leur production de liaisons, même s'ils maîtrisent les règles correspondantes.

Notre étude a un objectif plus général qui est l'enseignement de la prononciation des langues secondes et étrangères. Pour mieux enseigner, encore faut-il mieux comprendre les facteurs qui participent à l'acquisition d'une L1 et d'une L2, mais aussi mieux comprendre ce qui en empêche le processus.

Pour pouvoir mieux réaliser les liaisons et les enchaînements, les apprenants du français L2 ont besoin, dans un premier temps, de comprendre les règles principales de prononciation des liaisons et d'autres aspects phonétiques généraux. Car, si la compréhension ne suffit certes pas à la maîtrise des structures, du moins celle-ci aide à la mémorisation, incite les apprenants à faire attention à la forme et à comparer certains éléments avec leur L1, et répond aux questions éventuelles des apprenants qui appréhendent les questions phonétiques de façon analytique. L'explication n'est pas une fin en soit, mais n'est qu'un début, une prise de conscience. Les répétitions sont cruciales dans la maîtrise des structures. Les apprenants ont besoin d'entendre des structures et des phrases de nombreuses fois pour former leur oreille et adopter le nouveau système phonétique, et ce, idéalement, en début d'acquisition.

Des explications sans un nombre suffisant de répétitions auront pour effet que les apprenants comprennent et connaissent une règle, qu'ils seront peut-être capables de l'expliquer ou de répondre à des questions métalinguistiques, mais ne pourront pas appliquer

la règle de façon systématique. Pour qu'une règle devienne un automatisme, il faut que le cerveau ait créé des classes et soit capable de puiser dans une banque de données et dans les différentes catégories abstraites. Sans répétition et sans fréquence, les locuteurs n'ont pas assez d'items pour pouvoir créer des classes. S'ils n'ont pas créé de classes abstraites, ils ne peuvent qu'utiliser un nombre restreint d'items, les items les plus entendus, au coup par coup.

Les apprenants ont également besoin d'être capables de faire une analyse syntaxique correcte (Özçelik, 2008) pour déterminer dans quel cas on fait une liaison, mais aussi pour évaluer les priorités dans la hiérarchie des liaisons à réaliser. Comme le précise Grammont (1914), cité dans Laurent (2007 : 60) : « la difficulté est de savoir dans quels cas il faut lier et dans quels cas on doit s'en abstenir [...] »

Ils ont besoin de faire des exercices intensifs d'écoute, de discrimination auditive et de porter une attention particulière sur certains points qui leur sont problématiques, dans le cadre d'une analyse contrastive. En effet, comme l'ont montré Trahey et White (1993), les apprenants remarquent qu'une structure de la langue cible est différente de leur L1, mais ne reconnaissent pas une structure fautive utilisant les paramètres de leur L1. Cela indique que la fréquence seule est suffisante pour reconnaître la bonne formation de certaines structures, mais ne suffit souvent pas pour éviter les erreurs et notamment causées par le transfert de la L1. Il faut donc faire un travail particulier basé sur l'analyse contrastive.

Des exercices permettant une fréquence d'écoute et de production, avec de nombreuses répétitions et corrections auraient plusieurs fonctions : attirer l'attention des apprenants sur des éléments phonétiques spécifiques (ici les liaisons), leur fournir un nombre suffisant d'items permettant de créer des représentations abstraites, tout en travaillant les éléments particulièrement problématiques par rapport à la L2. Ainsi, un cours de prononciation permet d'attirer l'attention des apprenants sur des éléments souvent non indispensables à la communication, de leur fournir la fréquence de certaines structures et

cela, couplé à des présentations théoriques, permettrait de contrebalancer les effets d'interférence de la L1 et le manque d'input de la L2.

Une fois un inventaire plus riche mis en place, les apprenants devraient alors avoir constitué des représentations abstraites. A ce moment-là, ils pourront bénéficier de pratique intensive pour développer les représentations abstraites existantes et corriger les représentations incomplètes et fautives. Si les apprenants n'ont eu ni input riche, ni cours de prononciation en début d'acquisition, ils auront besoin de très nombreuses répétitions pour corriger les erreurs fossilisées. Ils auront aussi besoin de nombreuses corrections pendant le processus.

Après ces étapes de mise en place, ils feront des exercices de répétition et de lecture à haute voix de textes connus et inconnus et des exercices guidés de communication, les incitant à utiliser les structures étudiées, mais dans lesquels il ne faut pas beaucoup réfléchir au contenu des réponses. Par exemple, un étudiant pose une question : *Qu'est-ce que tu aimes regarder ?* (le but de l'exercice pourrait être d'utiliser un article pluriel + mot commençant par voyelle). Réponse : *des \_enfants qui jouent, des \_émissions de variété...*

Nous pensons que l'apprentissage de la prononciation devrait être au centre de l'enseignement. Il est possible d'être compris même avec des erreurs de grammaire mais il est difficile d'être compris si la prononciation et le vocabulaire sont trop éloignés de la langue cible. De plus, la production est également liée à d'autres composantes linguistiques. Elle est liée, entre autres, à la compréhension. Par exemple, si les apprenants ne font pas les enchaînements ou prononcent le schwa quand il ne faudrait pas le faire, ils ne comprennent pas non plus les locuteurs natifs si ceux-ci font des enchaînements, et ne prononcent pas le schwa (Saunders, 2007).

En outre, même si un accent très marqué n'entrave pas toujours la compréhension, elle distrait le locuteur, lui impose de faire des efforts pour s'adapter à un système phonétique

(souvent plus proche du système phonétique de la L1 du locuteur que de la langue cible) qui lui est étranger. Ces efforts sont fatigants pour le locuteur natif, lui demandant plus d'attention, ce qui au final peut décourager un locuteur d'écouter un apprenant. Quant aux locuteurs natifs non-habitués à écouter un accent étranger empruntant les caractéristiques d'un autre système phonétique, ils ne comprendront pas un apprenant à l'accent marqué.

De plus, un accent étranger marqué donne au locuteur natif l'impression que l'apprenant ne maîtrise pas bien la langue et peut même l'irriter (Tardif et d'Anglejan, 1981). Enfin, les erreurs de prononciation ont aussi des répercussions à l'écrit. Par exemple, si un apprenant prononce une consonne finale muette, telle que le t de *navet*, il y a de fortes chances pour qu'il écrive *navette* au lieu de *navet*.

Mais ce qui est encore plus important, ce sont les conséquences d'une prononciation fautive. Qu'on le veuille ou non, la langue des locuteurs L2 est jugée selon sa prononciation. Ainsi, un locuteur à la grammaire et au vocabulaire de niveau natif peut se voir refuser l'accès aux facultés d'éducation et aux postes d'enseignants ou de professeurs si sa prononciation est trop éloignée du français de référence. La prononciation est en cela encore plus importante que toutes les autres compétences linguistiques.

Nous avons analysé les effets de la fréquence et de l'enseignement des liaisons et enchaînements. Nous avons montré que la fréquence joue un rôle majeur dans l'acquisition des liaisons (et de la langue en général). Malheureusement toutes les universités et écoles et tous les cours de langue ne permettent pas toujours de fournir aux apprenants l'input dont ils auraient besoin. Il est par contre possible d'intégrer une partie phonétique aux cours de langues, et ceci est d'autant plus nécessaire que le nombre d'heures de cours est limité, pour compenser le manque d'input.

Nous avons vu que l'enseignement de la phonétique permet d'améliorer certains aspects de la prononciation des apprenants. Il est donc bénéfique de l'enseigner. Toutefois la

prononciation n'est souvent pas enseignée pour plusieurs raisons dont celles que décrit Jones (1997) :

Arguments against the explicit teaching of pronunciation rely on two basic assumptions about the acquisition of second language phonology: the first, based on the critical period hypothesis, claims that it is virtually impossible for adults to acquire native-like pronunciation in a foreign language; and the second, arising primarily from the work of Krashen (1982), insists that pronunciation is an acquired skill and that focused instruction is at best useless and at worst detrimental. (Jones, 1997 : 104)

De plus, les éléments que nous pensons être les plus utiles pour l'acquisition de la prononciation : la fréquence, l'écoute et les exercices de répétition sont souvent considérés par certains adeptes de l'approche communicative comme dépassés, inutiles, ennuyeux et peu efficaces, puisqu'ils ne rentrent pas dans le cadre de l'approche communicative, d'où l'abandon de tout ce qui peut rappeler des méthodes « anciennes », telles que la méthode audio-linguale, y compris les aspects positifs tels que les répétitions.

Lightbown et Spada (2006 : 184) résument assez bien (à notre avis) l'attitude de nombreux chercheurs et enseignants face au rôle de la répétition :

Some may find that they benefit from opportunities to imitate samples of the new language, and imitation is clearly important in developing pronunciation and intonation. For some advanced learners who are determined to improve their pronunciation, careful listening and imitation in a language laboratory can be very valuable. But for beginning learners, the slavish imitation and rote memorization that characterized audio-lingual language approaches to language teaching can lead to a dead end. Learners need to do more than recite bits of perfectly accurate language. (Lightbown and Spada, 2006 : 184)

L'approche communicative met, à juste titre, l'accent sur les activités de communication. Les apprenants de niveau B1-B2 parviennent (pour la plupart) à communiquer, à comprendre et se faire comprendre et en ce sens les objectifs ont été atteints. Les écoles d'immersion en sont la preuve vivante. Les élèves arrivent très vite à communiquer, et ce, plus rapidement qu'aucune autre approche. Toutefois, cette approche a aussi ses limites. Blanche (2004) considère que la prononciation n'est pas enseignée parce que, selon de nombreux enseignants, elle n'entrave pas la communication et n'est pas importante :

[...] FL teachers tend to view pronunciation as the least useful of the basic language skills, partly because it seems to be so much more difficult to improve than other skills. [...] yet the Communicative Approach does not stress the need for accent reduction – its real emphasis is on ‘intelligibility’, therefore on the ability to make oneself clearly understood despite a foreign accent. Consequently, pronunciation is still seen “as a small part of linguistic competence, which itself is a small part of communicative competence. (Blanche, 2004 : 176-178)

Pourtant, Lauret (2007 : 18) rappelle que bien prononcer, c’est d’abord être « intelligible ». L’intelligibilité dépend grandement des éléments suprasegmentaux. Or, l’entraînement de la prononciation est le plus souvent réduit au niveau segmental, alors que ce sont les éléments suprasegmentaux qui sont les plus importants pour l’intelligibilité (Lauret, 2007 : 24).

En effet, si un apprenant demande à un francophone, en parlant vite : « *Pyjama aux toilettes ?* » (à la place de « *Puis-je aller aux toilettes ?* ) ou si un francophone demande à un anglophone : « *Boite à musique, please ?* » (à la place de « *What time is it, please ?* ) en regardant sa montre, du moment que l’intonation est bonne, on les comprendra facilement en dépit des nombreux phonèmes fautifs!

De même, les enfants L1 sont le plus souvent compris de leur entourage, même s’ils ne prononcent pas correctement tous les phonèmes de leur L1 jusqu’à au moins 5 ou 6 ans. Et même dans ce cas, personne ne pensera qu’ils ont un accent « étrange » ou « étranger », car ils maîtrisent les éléments suprasegmentaux de leur L1. En outre, on peut être compris tout en ayant un accent étranger très fort, ce qui est généralement en rapport avec les éléments segmentaux (Lauret, 2007 : 23).

D’ailleurs, la question de l’intelligibilité est loin d’être clairement définie. Dans l’étude de Nagamine (2002), lorsque quatre locuteurs natifs ont dû donner leur avis sur les enregistrements de Japonais apprenant l’anglais, il s’est avéré que leurs interprétations n’étaient guère fiables, car les jugements n’étaient pas unanimes bien que les critères aient clairement été définis. L’intelligibilité, même avec des facteurs précis, n’est claire ni à définir ni à juger, même par les linguistes les plus expérimentés.



En effet, l'intelligibilité dépend de plusieurs facteurs, dont l'habitude qu'ont les locuteurs à entendre des accents étrangers, des connaissances qu'ont ces locuteurs de la L1 de l'apprenant, mais aussi de facteurs personnels tels que le degré d'acceptation des différences et de variabilité. Ainsi, un Français ne parlant pas l'anglais, aura sans doute plus de difficultés à comprendre un étudiant canadien anglophone parlant français avec un accent anglophone marqué qu'un québécois bilingue.

L'intelligibilité dépend aussi de la volonté de comprendre l'interlocuteur et de la situation et du contexte de communication. Bref, on ne peut absolument pas compter sur l'intelligibilité comme critère de prononciation, et ce, encore moins dans la salle de classe, où les apprenants sont généralement compris des professeurs comme des autres apprenants et la prononciation n'entrave qu'exceptionnellement la communication en salle de classe, puisque tout le monde comprend les erreurs (ici, tout le monde parle l'anglais).

Plus qu'une question d'intelligibilité, d'autres questions se posent quant aux raisons d'apprendre la prononciation d'une L2. Pourquoi restreindre les cours de langue à des exercices de « communication », ou des cours de grammaire et vocabulaire, quand la communication est en fait liée à de nombreux autres facteurs (tels que la précision de la prononciation) ? Pourquoi ne pas viser plus haut et avoir le meilleur des deux mondes : être capable de communiquer et être capable de parler de façon aussi proche que possible des locuteurs natifs ?

La prononciation native ou quasi-native est possible, si des cours d'orthoépée sont enseignés de façon régulière et appropriée aux besoins des apprenants. Notre étude a permis de montrer que les apprenants peuvent améliorer leur production de liaisons obligatoires de 10% après seulement 30 mn d'enseignement de la phonétique. Nous pouvons raisonnablement supposer qu'un enseignement d'au moins une heure par semaine pendant plusieurs années devrait permettre de grandement améliorer la prononciation des apprenants

## **Annexes**

## Annexe 1 : Directives du ministère d'Ontario

<http://www.edu.gov.on.ca/eng/curriculum/secondary/>

### Core French, Grade 11

Students should recognize and use these language structures in all three strands.

#### Nouns and Pronouns

##### Verbs

- expressions of quantity + *de* + a noun (e.g., *une paire de bottes, une douzaine de chapeaux, beaucoup d'hommes, peu d'enfants*)
- use of all relative pronouns, including *dont, ce dont, and où*
- demonstrative pronouns (*celui, celle, ceux, celles*)
- position of two object pronouns with imperatives (e.g., *Envoyez-les-moi. Ne m'en faites pas parvenir.*)
- formation of the *plus-que-parfait, futur antérieur, and conditionnel passé* of *-er, -ir, -re* verbs and irregular verbs
- sequence of tenses with *si* using the *plus-que parfait* and the *conditionnel passé* (e.g., *Si j'étais allé à New York, je serais allé voir une pièce de théâtre.*)
- sequence of present and past tenses used in indirect discourse (e.g., *Le premier ministre a dit qu'il allait parler de ce sujet dans un communiqué de presse. On nous a dit que l'explosion avait eu lieu à 15 h.*)
- formation and use of the *participe présent* (e.g., *En me promenant dans le quartier, j'ai vu une mouffette dans le garage d'un voisin.*)
- recognition of the *passé simple* as the literary tense equivalent to the *passé composé*
- expressions of quantity + *de* + a noun (e.g., *une paire de bottes, une douzaine de chapeaux, beaucoup d'hommes, peu d'enfants*)
- use of all relative pronouns, including *dont, ce dont, and où*
- demonstrative pronouns (*celui, celle, ceux, celles*)
- sequence of tenses with *si* using the *imparfait* and the *conditionnel présent* (e.g., *Si j'étais malade, j'irais chez le médecin.*)
- use of the *subjonctif présent* with impersonal expressions (e.g., *il faut, il est important, il est essentiel*) (e.g., *Il faut que les élèves fassent leurs devoirs.*)
- *passé composé* of reflexive and pronominal verbs (e.g., *Je me suis habillé. Nous nous sommes rencontrés au centre commercial.*)
- recognition of the *passé simple* as the literary tense equivalent to the *passé composé*

### Core French, Grade 12, University Preparation (FSF4U)

This course draws on a variety of themes to promote extensive development of French language skills. Students will consolidate their oral skills as they discuss literature, culture, and current issues. They will read a variety of texts and will write a formal essay. The use of correct grammar and appropriate language conventions in both spoken and written French will be emphasized throughout the course.

**Prerequisite:** Core French, Grade 11, University Preparation

THE ONTARIO CURRICULUM, GRADES 11 AND 12: FRENCH AS A SECOND LANGUAGE

Oral Communication

#### Overall Expectations

By the end of this course, students will:

- respond in a variety of ways to a wide range of media works;
- express and justify ideas and opinions in self-directed conversations and discussions on a variety of issues;
- make oral presentations on a variety of topics;
- use appropriate language conventions during oral communication activities.

#### Specific Expectations

##### Listening

By the end of this course, students will:

- summarize the content of a variety of media works (e.g., full-length films, recorded debates, songs, television and radio programs);
- demonstrate an understanding of oral presentations by classmates and other speakers (e.g., by relating the content to other contexts, discussing alternative viewpoints, predicting future trends);
- listen to and demonstrate an understanding of text spoken in a variety of French accents (by restating the main idea, identifying specific information, agreeing or disagreeing with the author's point of view).

### *Speaking*

By the end of this course, students will:

- express, support, and explain their viewpoints in formal debates related to topics under study;
- give a researched presentation (minimum 5 minutes in length) supported by point form notes and visual materials, and answer questions posed by classmates;
- use critical-thinking skills in responding to classmates' presentations (e.g., analyse the information, develop a counter-argument);
- dramatize a scene from a play to demonstrate an understanding of its plot and characterization;
- analyse an issue under discussion and recommend possible solutions.

### *Application of Language Conventions*

By the end of this course, students will:

- recognize and use appropriate language structures (see language structures for Core French, Grade 12, p. 25);
- use newly acquired vocabulary in conversation;
- use stylistic devices (e.g., juxtaposition, similes, metaphors, hyperbole) to add interest to their spoken French.

### Oral Communication

#### **Overall Expectations**

By the end of this course, students will:

- respond in a variety of ways to a range of media works;
- express and justify ideas and opinions in self-directed conversations and discussions on a variety of business-related issues;
- make oral presentations on a variety of topics;
- use correct grammar and appropriate language conventions during oral communication activities.

#### **Specific Expectations**

##### *Listening*

By the end of this course, students will:

- understand and participate in spontaneous conversations (e.g., when working with classmates on a collaborative project);
- take point-form notes while listening to discussions, news reports, and short stories, and refer to them in oral or written assignments;
- ask questions and make comments to clarify, evaluate, and provide feedback on classmates' presentations.

##### *Speaking*

By the end of this course, students will:

- use critical-thinking skills in responding to classmates' presentations (e.g., analyse the information, elaborate on the ideas presented);
- role-play situations (e.g., seeking information, resolving conflicts between people) based on a topic under study;
- contribute ideas in class or small-group discussions (e.g., suggest a different point of view, critique an argument, relate the topic to a similar issue in another context);
- give a researched presentation on a business topic (e.g., an analysis of the coverage of a business news story, a review of careers where knowledge of French is an asset);
- conduct a mock interview, referring to information contained in a prepared résumé.

### *Application of Language Conventions*

By the end of this course, students will:

- recognize and use appropriate language structures (see language structures for Core French, Grade 12, p. 25);
- use newly acquired vocabulary in conversation.

### Language Structures

#### **Core French, Grade 12**

Students should recognize and use these language structures in all three strands.

#### **Nouns and Pronouns**

##### **Verbs**

##### **Prepositions and**

##### **Conjunctions**

##### **Negation**

- possessive pronouns (e.g., *le mien, la tienne, le vôtre, les nôtres*)

- interrogative pronoun *lequel* (e.g., *Lequel des deux garçons est votre fils?*)
- use of the pronoun *on* to reflect the English passive (e.g., *Ici on parle français.* French is spoken here.)
- relative pronoun *lequel* (e.g., *Voilà la raison pour laquelle il est parti si vite.*)
- tense sequence in past narration (*passé composé*, *imparfait*, *plus-que-parfait*) (e.g., *Quand je suis arrivé chez mon ami, il regardait le match de football. Il avait déjà mis les amuse-gueule devant la télé.*)
- use of the *subjonctif présent* after conjunctions (e.g., *sans que*, *de peur que*, *à moins que*)
- verbs of perception (e.g., *écouter*, *entendre*, *regarder*) and *laisser* followed by the infinitive (e.g., *Je regarde jouer les enfants. Je les laisse jouer dans la cour.*)
- sequence of present and past tenses with *les expressions temporelles* (e.g., *depuis*, *ça fait*, *il y a*) (e.g., *Il habite ici depuis deux ans. Ça fait deux ans qu'il habite ici. Il y a deux ans qu'il est parti de sa ville natale.*)
- formation and use of the *subjonctif passé* of *-er*, *-ir*, and *-re* verbs and irregular verbs
- formation and use of the *infinitif passé* (e.g., *Après avoir regardé les nouvelles, je me suis couché.*)
- *faire causatif* (e.g., *Le prof fait toujours visionner les films dans la grande salle.*)
- use of the *subjonctif présent* after expressions of emotion, wish, order, permission (e.g., *Je suis content qu'il puisse venir. Il veut que je vienne aussi. J'aimerais que mes devoirs soient faciles.*)
- *peut-être* vs. *peut-être que* (e.g., *Va-t-il venir? Peut-être.* vs. *Peut-être qu'il va venir.*)
- use and position of *ne . . . jamais*, *ne . . . rien*, *ne . . . nulle part*, *ne . . . personne*, *ne . . . plus*, *ne . . . aucun* (e.g., *Je n'ai jamais lu ce rapport. Personne n'est venu me rendre visite quand j'étais malade.*)
- negative infinitives (e.g., *C'est un film à ne pas manquer.*)
- possessive pronouns (e.g., *le mien*, *la tienne*, *le vôtre*, *les nôtres*)
- interrogative pronoun *lequel* (e.g., *Lequel des deux garçons est votre fils?*)
- formation of the *plus-que-parfait*, *futur antérieur*, and *conditionnel passé* of *-er*, *-ir*, *-re* verbs and irregular verbs
- formation and use of the *participe présent* (e.g., *En me promenant dans le quartier, j'ai vu une mouffette dans le garage d'un voisin.*)
- sequence of present and past tenses with *les expressions temporelles* (e.g., *depuis*, *ça fait*, *il y a*) (e.g., *Il habite ici depuis deux ans. Ça fait deux ans qu'il habite ici. Il y a deux ans qu'il est parti de sa ville natale.*)
- use of the *subjonctif présent* after expressions of emotion, wish, order, permission (e.g., *Je suis content qu'il puisse venir. Il veut que je vienne aussi. J'aimerais que mes devoirs soient faciles.*)
- *peut-être* vs. *peut-être que* (e.g., *Va-t-il venir? Peut-être.* vs. *Peut-être qu'il va venir.*)
- use and position of *ne . . . jamais*, *ne . . . rien*, *ne . . . nulle part*, *ne . . . personne*, *ne . . . plus*, *ne . . . aucun* (e.g., *Je n'ai jamais lu ce rapport. Personne n'est venu me rendre visite quand j'étais malade.*)
- negative infinitives (e.g., *C'est un film à ne pas manquer.*)

### **Extended French, Grade 11, University Preparation (FEF3U)**

This course focuses on developing French-language skills through the study of Canadian francophone authors. Students will analyse a range of works and produce written assignments in a variety of genres, including the formal essay. The use of correct grammar and appropriate language conventions in both spoken and written French will be emphasized throughout the course.

**Prerequisite:** Extended French, Grade 10, Academic

EXTENDED FRENCH, GRADE 11, UNIVERSITY PREPARATION (FEF3U) 27

Oral Communication

#### **Overall Expectations**

By the end of this course, students will:

- respond in a variety of ways to a wide range of spoken texts and media works;
- discuss and debate topics based on class discussions, individual research, and personal interests;
- use correct grammar and appropriate language conventions during oral communication activities.

#### **Specific Expectations**

##### *Listening*

By the end of this course, students will:

- follow a series of complicated live or recorded instructions (e.g., how to prepare a complete meal);
- take point-form notes during a classmate's or a guest's presentation, and refer to these notes in participating in a question-and-answer discussion about the presentation;
- summarize and answer questions in response to media works (e.g., videos, radio broadcasts) and explain their points of view.

### *Speaking*

By the end of this course, students will:

- express their ideas and opinions about a media work or a written text (e.g., a film, a magazine article) and give reasons for their preferences or point of view;
- give a presentation based on research and respond appropriately to questions from the audience;
- express themselves clearly and coherently in impromptu discussions on topics under study;
- prepare and conduct a survey on a topic under study, compile results, and present their findings in a report;
- participate in a panel discussion by contributing ideas, supporting interpretations and viewpoints, and elaborating on and questioning the ideas of others.

### *Application of Language Conventions*

By the end of this course, students will:

- recognize and use appropriate language structures (see language structures for Extended French, Grade 11, p. 34);
- use newly acquired vocabulary in conversations, discussions, and presentations;
- identify and correct anglicisms and errors in their speech;
- incorporate colloquialisms and idiomatic expressions into their speech;
- select vocabulary and language structures to enhance the clarity and precision of their speech. (p.25-27)

## EXTENDED FRENCH, GRADE 12, UNIVERSITY PREPARATION (FEF4U) 31

### Oral Communication

#### **Overall Expectations**

By the end of this course, students will:

- respond to a wide range of spoken texts and media works;
- discuss and debate topics based on class discussions, individual research, and personal interests;
- use correct grammar and appropriate language conventions during oral communication activities.

#### **Specific Expectations**

##### *Listening*

By the end of this course, students will:

- understand and explain different points of view made during an interview or a debate;
- demonstrate an understanding of media works by expressing ideas and opinions about points of view and issues discussed;
- identify the inferences made during radio or television broadcasts;
- demonstrate the ability to detect nuances of language in various forms of oral communication.

##### *Speaking*

By the end of this course, students will:

- present orally an independently researched topic and respond to questions from the audience;
- summarize in sequence the essential points of a narration or discussion;
- use facts to support or refute arguments or opinions;
- provide spontaneous and appropriate answers to questions on a variety of topics;
- participate in large- and small-group discussions, round-tables, and debates relating to a particular theme by explaining and justifying their points of view and reacting to those of their classmates;
- lead a group discussion on a particular topic by encouraging individual participation, inviting comments from group members, and asking questions to stimulate the discussion.

##### *Application of Language Conventions*

By the end of this course, students will:

- recognize and use appropriate language structures (see language structures for Extended French, Grade 12, p. 34);
- use newly acquired vocabulary in conversations, discussions, and presentations;
- identify and correct anglicisms and errors in their speech;
- incorporate colloquialisms and idiomatic expressions into their speech;
- select vocabulary and language structures to enhance the clarity and precision of their speech.

Oral Communication

**Overall Expectations**

By the end of this course, students will:

- understand the content and interpret the message conveyed by a variety of media works;
- speak with native-like fluency on a variety of topics ranging from personal life to global issues;
- express and explain abstract ideas;
- use correct grammar and appropriate language conventions during oral communication activities.

**Specific Expectations**

*Listening*

By the end of this course, students will:

- demonstrate an understanding of media works by expressing ideas and opinions about the points of view and issues discussed;
- demonstrate an understanding of complex, informal conversations delivered with normal clarity and speed (e.g., in *québécois*, in French films);
- identify and explain errors and inconsistencies in persuasive speech (e.g., contradictions, factual errors, and false generalizations in media reports or debates).

*Speaking*

By the end of this course, students will:

- express clearly and confidently their personal point of view in informal discussions;
- deliver a well-organized, well-thought-out presentation clearly and with confidence;
- debate formally and informally issues arising from reading selections;
- expand upon issues arising from viewing or listening to media works by analysing the message, expressing their point of view, and suggesting possible solutions.

*Application of Language Conventions*

By the end of this course, students will:

- recognize and use appropriate language structures (see language structures for French Immersion, Grade 12, p. 43);
- use newly acquired vocabulary in conversations, discussions, and presentations;
- identify and correct anglicisms and errors in their speech;
- incorporate colloquialisms and idiomatic expressions into their speech;
- select vocabulary and language structures to enhance the clarity and precision of their speech.

**French Immersion, Grade 12, University Preparation (FIF4U) p. 40**

Language Structures

**French Immersion, Grades 11 and 12**

Students should recognize and use these language structures in all three strands.

**Nouns and Pronouns**

**Verbs<sup>1</sup>**

- indefinite pronouns *plusieurs, chacun, personne, aucun, rien*
- *participe passé* used as an adjective (e.g., *Arrivés à Paris, ils sont allés au restaurant.*)
- use of the *participe passé composé* to describe an action that takes place before another one (e.g., *Ayant pris l’avion, ils sont vite arrivés à Montréal.*)
- verbs of perception (e.g., *écouter, entendre, regarder*) and *laisser* followed by the infinitive (e.g., *Je regarde jouer les enfants. Je les laisse jouer dans la cour.*)
- formation and use of the *subjonctif passé* of *-er*, *-ir*, and *-re* verbs and irregular verbs
- use of the *subjonctif présent* and *subjonctif passé* after conjunctions (e.g., *sans que, de peur que, à moins que*)
- recognition of the *passé simple* as the literary tense equivalent to the *passé composé* **12**
- use of the *subjonctif* after indefinite expressions (e.g., *Je cherche quelqu’un qui sache la réponse.*)
- use of *on* and *se* to avoid the passive (e.g., *On construira la maison l’année prochaine. La porte se fermera à minuit.*)

*Note:* The Grade 12 French Immersion course will consolidate all language structures studied throughout the French Immersion program (p.43)

## **Communication**

- communication of information and ideas
- use of language
- communication for different audiences and purposes, using various forms

### **50–59% (Level 1)**

- demonstrates limited knowledge of language forms and conventions
- demonstrates limited understanding of content
- uses critical and creative thinking skills with limited effectiveness
- applies few of the skills involved in an inquiry process
- communicates information and ideas with limited clarity
- uses language with limited accuracy and effectiveness
- communicates with a limited sense of audience and purpose, using few appropriate forms

### **60–69% (Level 2)**

- demonstrates some knowledge of language forms and conventions
- demonstrates some understanding of content
- uses critical and creative thinking skills with moderate effectiveness
- applies some of the skills involved in an inquiry process
- communicates information and ideas with some clarity
- uses language with some accuracy and effectiveness
- communicates with some sense of audience and purpose, using some appropriate forms

### **70–79% (Level 3)**

- demonstrates considerable knowledge of language forms and conventions
- demonstrates considerable understanding of content
- uses critical and creative thinking skills with considerable effectiveness
- applies most of the skills involved in an inquiry process
- communicates information and ideas with considerable clarity
- uses language with considerable accuracy and effectiveness
- communicates with a clear sense of audience and purpose, using appropriate forms

### **80–100% (Level 4)**

- demonstrates thorough knowledge of language forms and conventions
- demonstrates thorough understanding of content
- uses critical and creative thinking skills with a high degree of effectiveness
- applies all or almost all of the skills involved in an inquiry process
- communicates information and ideas with a high degree of clarity, and with confidence
- uses language with a high degree of accuracy and effectiveness
- communicates with a strong sense of audience and purpose, using appropriate forms



## **Annexe 2 : description<sup>45</sup> du Cadre Européen Commun de Référence (CECR)**

C2 Comme C1

C1 Peut varier l'intonation et placer l'accent phrastique correctement afin d'exprimer de fines nuances de sens.

B2 A acquis une prononciation et une intonation claires et naturelles.

B1 La prononciation est clairement intelligible même si un accent étranger est quelquefois perceptible et si des erreurs de prononciation proviennent occasionnellement.

A2 La prononciation est en général suffisamment claire pour être comprise malgré un net accent étranger mais l'interlocuteur devra parfois faire répéter.

A1 La prononciation d'un répertoire très limité d'expressions et de mots mémorisés est compréhensible avec quelque effort pour un locuteur natif habitué aux locuteurs du groupe linguistique de l'apprenant/utilisateur.

---

<sup>45</sup> [http://www.coe.int/t/dg4/linguistic/Source/Framework\\_FR.pdf](http://www.coe.int/t/dg4/linguistic/Source/Framework_FR.pdf) p.92

### Annexe 3 : questionnaire pour les participants anglophones

#### 1. Renseignements personnels

1. Nom, prénom: \_\_\_\_\_
2. Adresse: \_\_\_\_\_
3. n° de téléphone: \_\_\_\_\_
4. e-mail \_\_\_\_\_
5. Sexe: Masculin Féminin
6. Age: \_\_\_\_\_
7. Pays et ville de naissance: \_\_\_\_\_
8. Nationalité: \_\_\_\_\_
9. Langue(s) maternelle(s): \_\_\_\_\_
10. Autres langues parlées: \_\_\_\_\_

#### 2. Formation scolaire

1. Pendant combien d'années avez-vous étudié le français ? \_\_\_\_\_  
(sans compter cette année) **NOMBRE TOTAL D'ANNÉES**

2. Dans quel système (immersion, « core French », « extended French », francophone...) ?  
\_\_\_\_\_

3. Combien d'heures par semaine avez-vous étudié/parlé le français chaque semaine ?  
Exemple : de 6 à 8 ans immersion 30h /semaine

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

4. Dans quel système avez-vous étudié et pendant combien d'années ?

	1. core French	2. extended French	3. immersion	4. francophone	5. autre
Type d'école					
Nombre d'années					
heures /semaine					
Age					

5. Nom de l'école ou des écoles \_\_\_\_\_

6. Avez-vous passé du temps en région francophone quand vous étiez à l'école ? oui non  
Si oui, pendant combien de temps ? Quel âge aviez-vous ? \_\_\_\_\_

### 3. Compétences langagières (self ratings) en français

1. Quelle note finale pensez-vous recevoir dans ce cours ?

-60    60-65    65-70    70-75    75-80    80-85    85-90    90-95    95+

2. À quel niveau estimez-vous **vosre prononciation**, comparée à un locuteur natif

Ex : 60= très éloignée d'un natif                      80= assez proche                      95+ comme un natif

-60    60-65    65-70    70-75    75-80    80-85    85-90    90-95    95+

3. À quel niveau estimez-vous **vosre vocabulaire**, comparé à un locuteur natif

Ex : 60= très éloigné d'un natif                      80= assez proche                      95+ comme un natif

-60    60-65    65-70    70-75    75-80    80-85    85-90    90-95    95+

4. À quel niveau estimez-vous **vosre grammaire**, comparée à un locuteur natif

Ex : 60= très éloignée d'un natif                      80= assez proche                      95+ comme un natif

-60    60-65    65-70    70-75    75-80    80-85    85-90    90-95    95+

5. À quel niveau estimez-vous **vosre capacité à communiquer**, comparée à un locuteur natif

Ex : 60= très éloignée d'un natif                      80= assez proche                      95+ comme un natif

-60    60-65    65-70    70-75    75-80    80-85    85-90    90-95    95+

### 4. Questions générales d'attitude face à la prononciation

Veillez répondre aux questions suivantes en utilisant les chiffres ci-dessous

5 = c'est **toujours ou presque toujours vrai** pour moi

4 = c'est **souvent vrai** pour moi

3 = c'est **quelquefois vrai** pour moi

2 = **d'habitude ce n'est pas vrai** pour moi

1 = **ce n'est jamais jamais vrai** pour moi

\_\_\_\_\_ 1. Je veux avoir une bonne prononciation pour qu'on me comprenne bien.

\_\_\_\_\_ 2. Je fais attention à la prononciation du professeur ou/et des francophones.

\_\_\_\_\_ 3. Je m'entraîne souvent à répéter des mots ou phrases à la maison pour améliorer ma prononciation.

\_\_\_\_\_ 4. J'écoute avec attention les remarques du professeur sur ma prononciation, et j'essaie de ne plus refaire les mêmes erreurs.

\_\_\_\_\_ 5. Je demande de l'aide pour améliorer ma prononciation (demande au professeur ou à des francophones comment un mot se prononce, pose des questions sur ce que j'ai remarqué, etc.).

\_\_\_\_\_ 6. Je trouve des « trucs » qui m'aident à améliorer ma prononciation.

\_\_\_\_\_ 7. Je veux parler comme un francophone et je fais tout pour y arriver.

- \_\_\_\_\_ 8. Je crois que je peux améliorer ma prononciation si je reçois l'aide nécessaire.
- \_\_\_\_\_ 9. Je suis un bon imitateur/une bonne imitatrice. (good mimick).
- \_\_\_\_\_ 10. J'aime lire à haute voix en français.
- \_\_\_\_\_ 11. J'ai une oreille musicale. Je chante juste et j'entends quand on fait une fausse note.
- \_\_\_\_\_ 12. Quand je parle français, je me concentre sur la prononciation.
- \_\_\_\_\_ 13. Je suis capable d'expliquer des règles de prononciation en français.
- \_\_\_\_\_ 14. J'aimerais suivre un ou plusieurs cours de prononciation.
- \_\_\_\_\_ 15. J'aimerais qu'il y ait une partie prononciation à chaque cours de français.
- \_\_\_\_\_ 16. J'ai peur de parler en français au professeur ou en classe.
- \_\_\_\_\_ 17. J'ai honte de parler en français parce que je sais que ma prononciation n'est pas bonne.
- \_\_\_\_\_ 18. Je ne fais rien pour améliorer ma prononciation, elle s'améliorera toute seule.
- \_\_\_\_\_ 19. Je n'entends pas ou pas beaucoup de différences entre ma prononciation et celle des natifs
- \_\_\_\_\_ 20. J'entends de grandes différences entre ma prononciation et celle des natifs, mais je ne sais pas exactement quels sont mes problèmes de prononciation.

### 5. Faites-vous l'une de ces activités en français, en dehors de vos cours ?

Si vous répondez « oui », combien de fois par mois ?

- |  |     |     |       |
|--|-----|-----|-------|
| 1. parler français avec de la famille ou des amis Francophones ? | oui | non | _____ |
| 2. parler français dans un club de français                      | oui | non | _____ |
| 3. regarder la télévision ou voir des films au cinéma            | oui | non | _____ |
| 4. écouter des chansons  | oui | non | _____ |
| 5. apprendre des chansons francophones par cœur                  | oui | non | _____ |
| 6. écouter la radio  | oui | non | _____ |
| 7. lire  | oui | non | _____ |
| 8. consulter l'internet  | oui | non | _____ |
| 9. écrire des e-mails en français                                | oui | non | _____ |
| 10. autre activité en français                                   | oui | non | _____ |

### 6. Questions sur la prononciation

1. Avez-vous appris des règles de prononciation dans vos cours de langue ? oui – non  
Si oui, des règles portant sur quels domaines ? Entourez les éléments étudiés :

Intonation    accentuation de syllabe    prononciation de sons    liaisons    autres

2. Vos professeurs corrigeaient-ils votre prononciation ? oui – non  
Si oui, quel âge aviez-vous ? \_\_\_\_\_

Si oui, quels aspects de votre prononciation ?

	toujours	très souvent	souvent	parfois	rarement	jamais
prononciation de sons isolés (u, r)						
intonation						
accentuation						
E muet						
liaisons						
autre						
De quel âge à quel âge ?						

3. Comment les professeurs vous corrigeaient-ils ?

1. En vous donnant la bonne réponse que vous deviez répéter ? oui – non
2. En vous demandant de répéter et vous deviez trouver vous-mêmes? oui – non

D'une autre façon ? Si oui, expliquez.

4. Si vous pensez que la prononciation est importante, expliquez pourquoi. Expliquez vos expériences avec la prononciation (vos sentiments, ce qui a été fait en cours, etc.)

### 7. Questions spécifiques sur les liaisons

1. Pouvez-vous donner un exemple d'une liaison obligatoire (qu'on doit toujours faire) ?

2. Pouvez-vous donner un exemple d'une liaison facultative (qu'on fait si on veut) ?

3. Vous paraît-il important de faire les liaisons, oui, non, pourquoi ?

4. Que savez-vous sur le sujet des liaisons? Connaissez-vous des règles ? Si oui, lesquelles ?  
Par exemple : les \_enfants sont \_arrivés

**Annexe 4 : Questionnaire sur ce que les étudiants anglophones ayant participé à l'étude ont pensé de la leçon de phonétique (distribué au dernier cours)**

1. Comment avez-vous trouvé cette petite présentation pour améliorer votre prononciation?

1. inutile
2. assez utile
3. utile
4. très utile

2. Comment avez-vous trouvé cette petite présentation pour améliorer votre compréhension et perception du français parlé par des francophones ?

1. inutile
2. assez utile
3. utile
4. très utile

3. Comment les concepts utilisés pour cette présentation vous ont-ils semblé?

1. très difficiles à comprendre
2. assez difficile à comprendre
3. difficiles à comprendre
4. faciles à comprendre
5. très faciles à comprendre

4. Comment pensez-vous vous être amélioré dans le domaine des liaisons, compte tenu du temps limité ?

Je pense que

1. J'ai fait beaucoup de progrès
2. J'ai fait quelques progrès
3. Je n'ai pas fait de progrès
4. J'ai régressé

5. Commentaires généraux. Qu'avez-vous pensé de cette présentation ? Qu'est-ce qui vous a semblé le moins utile ou le plus utile ?

## **Annexe 5 : questionnaire pour les participants francophones**

### **1. Renseignements personnels**

1. Nom, prénom: \_\_\_\_\_
2. Adresse, n° de téléphone: \_\_\_\_\_
3. Sexe: Masculin Féminin
4. Age: \_\_\_\_\_
5. Pays et ville de naissance: \_\_\_\_\_
6. Nationalité: \_\_\_\_\_
7. Langue(s) maternelle(s): \_\_\_\_\_
8. Autres langues parlées: \_\_\_\_\_
9. Plus haut diplôme obtenu : \_\_\_\_\_
10. Matière(s) de spécialisation : \_\_\_\_\_

### **2. Questions la prononciation**

1. A l'école primaire et secondaire, vos professeurs corrigeaient-ils votre prononciation quand vous parliez ? oui – non

Si oui, quel âge aviez-vous ? \_\_\_\_\_

2. avez-vous fait de la lecture à haute voix pendant laquelle un enseignant corrigeait votre prononciation si vous faisiez des fautes ? oui – non

Si oui, expliquez.

### **3. Questions spécifiques sur les liaisons**

1. Pouvez-vous donner un exemple d'une liaison obligatoire (qu'on doit toujours faire) ?

2. Pouvez-vous donner un exemple d'une liaison facultative (qu'on fait si on veut) ?

3. Vous paraît-il important de faire les liaisons, oui, non pourquoi ?

4. Que savez-vous sur le sujet des liaisons? Connaissez-vous des règles ? Si oui, lesquelles ?  
Par exemple : les\_enfants sont\_arrivés

5. Autres commentaires en rapport avec ce sujet.

## Annexe 6 : photocopies distribuées et expliquées en classe – Cours sur les liaisons

Définition du Petit Robert 2002 (p.1482). **Liaison** : action de prononcer deux mots consécutifs en unissant la consonne finale du premier mot (non prononcée isolément ou devant une consonne) à la voyelle initiale du mot suivant (ex. les petits\_enfants).

### I. LIAISONS OBLIGATOIRES DANS LE GROUPE NOMINAL

#### 1. Articles + nom (ou adjectif) commençant par une voyelle

UN :	un (n)angle un (n)œuf	un (n)avion un (n)ours	un (n)aéroport un (n)ordinateur	un (n)aigle un (n)exposé	un (n)éléphant un (n)alphabet
LES :	les alibis les (z)aérosols les (z)amandes	les (z)alarmes les (z)amoureux les (z)animaux	les (z)accents les (z)années	les (z)Allemands les (z)autorités	
DES	des (z)Anglais des (z)analyses des (z)amours	des (z)Américains des (z)amateurs des (z)allergies	des (z)Espagnols des (z)amis	des (z)Italiens des (z)ambulances	
AUX	aux (z)anges aux (z)armes aux (z)autorités	aux (z)antipodes aux (z)alentours aux (z)extraits	aux (z)épinards aux (z)épaules	aux (z)archives aux (z)urgences	

#### 2. Articles + nom (ou adjectif) commençant par un H muet

La plupart des mots qui commencent par un H ont un H muet. Cela veut dire qu'on **FAIT l'élision** (on met une apostrophe, on enlève le e de *le* et le a de *la*) ET on **FAIT la liaison**.

L'homme, l'hôpital, un habit, l'huitre, l'hippopotame, l'habitude, l'habitant,  
Des (z)hommes, les (z)hôpitaux, des (z)habits, des (z)huitres, des (z)hippopotames, des (z)habitudes, des (z)habitants.

#### 3. Adjectifs + nom

##### • Adjectifs possessifs

MES	mes (z)opinions, mes (z)idées, mes (z)aventures,
TES	tes (z)enfants, tes (z)amis, tes (z)oncles
SES	ses (z)ennuis, ses (z)examens, ses (z)impressions
NOS	nos (z)amis, nos (z)idées, nos (z)enfants
VOS	vos (z)étudiants, vos (z)excuses, vos (z)études
LEURS	leurs (z)enfants, leurs (z)amis, leurs (z)excuses



• **Adjectifs avec nasales**

MON : mon (N)instrument, mon (N)exploit, mon (N)écriture

TON : ton (N)emploi, ton (N)épaule, ton (N)avion

SON : son (N)idée, son (N)entrée, son (N)ulcère

BON : bon (N)étudiant, bon (N)avocat, bon (N)état

MOYEN : Moyen-(N)Orient, moyen-(N)âge

CERTAIN : certain (N)âge, certain (N)effort, certain(N) homme

PLEIN : plein (N)emploi, plein (N)air, plein (N)effort

AUCUN : aucun (N)avantage, aucun (N)effort, aucun (N)endroit

• **Adjectifs de BAGS (Beauty, Age, Goodness, Size): courts, les plus fréquents, se plaçant TOUJOURS AVANT le nom**

adjectif	Masc. Sing.	Masc. plur.	Fém. plur.
ancien	<b>ancien (n)immeuble</b> <b>ancien (n)agenda</b>	<b>anciens (z)amis</b> <b>anciens (z)uniformes</b>	anciennes (z)amies anciennes (z)installations
aucun	<b>aucun (n)effet</b> <b>aucun (n)espoir</b>	pluriel rare	pluriel rare
bon	<b>bon (n)appétit</b> <b>bon (n)été</b>	<b>bons (z)assistants</b> <b>bons (z)étudiants</b>	<b>bonnes (z)excuses</b> <b>bonnes (z)idées</b>
beau	(bel : enchaînement)	<b>beaux (z)enfant</b> <b>beaux-(z)arts</b>	belles (z)araignées belles (z)images
faux	<b>faux-(z)air</b> <b>faux-(z)ami</b>	<b>faux (z)airs</b> <b>faux-(z)amis</b>	fausses (z)illusions fausses (z)économies
fou	(fol : enchaînement)	<b>fous (z)architectes</b> <b>fous (z)espoirs</b>	folles (z)aventures folles (z)aventures
grand	<b>grand (t)immeuble</b> <b>grand (t)angle</b>	<b>grands (z)arbres</b> <b>grands (z)enfants</b>	grandes (z)idées grandes (z)orgues
gros	<b>un gros (z)appétit</b> <b>un gros (z)effort</b>	<b>de gros (z)ennuis</b> <b>de gros (z)yeux</b>	grosses (z)assiettes grosses (z)oreilles
jeune	(jeune : enchaînement)	jeunes (z)espoirs jeunes (z)enfants	jeunes (z)antilopes jeunes (z)adolescentes
joli	Pas de consonne de liaison	<b>jolis (z)hivers</b> <b>jolis (z)endroits</b>	jolies (z)histoires jolies (z)italiennes
long	(long= liaison rare)	<b>longs (z)hivers</b> <b>longs (z)étés</b>	longues (z)antennes longues (z)observations
meilleur	(meilleur : enchaînement))	meilleurs (z)amis meilleurs (z)avocats	meilleures (z)armes meilleures (z)étudiantes
mauvais	<b>mauvais (z)état</b> <b>mauvais (z)exemple</b>	<b>mauvais (z)exemples</b> <b>mauvais (z)esprit</b>	mauvaises (z)intentions mauvaises (z)impressions
nouveau	(nouvel : enchaînement))	<b>nouveaux(z) amis</b> <b>nouveaux (z)exploits</b>	nouvelles (z)odeurs nouvelles (z)idées
petit	<b>petit (t)homme</b> <b>petit (t)ami</b>	<b>petits (z)efforts</b> <b>petits (z)ennuis</b>	petites (z)annonces petites (z)inventions
vieux	(vieil : enchaînement))	<b>vieux (z)apaches</b> <b>vieux (z)écrits</b>	vieilles (z)affaires vieilles (z)habitudes

• **Adjectifs avec consonne latente se prononçant comme au féminin**

Premier : premier (R)amour, premier (R)effort, premier (R)étage, premier (R)avril

Dernier : dernier (R)étage, dernier (R)accident, dernier (R)homme

Léger : léger (R)ennui, léger (R)accent, léger (R)effort

Ancien : l'ancien (N)édifice, l'ancien (N)élève, ancien (N)immeuble

• **Adjectifs dont la consonne change** d se prononce T en liaison

Second (T)étage, second (T)accident,

Un grand (T)appartement, un grand (T)enfant, un grand (T)immeuble

• **Adjectifs cardinaux « réguliers »**

Deux (Z)amis, trois (Z)amis, six (Z)amis, dix (Z)amis

Deux (Z)ans, trois (Z)ans, six (Z)ans, dix (Z)ans

Six (Z)heures, six (Z)ans, six (Z)enfants

Dix (Z)arbres, dix(Z) heures, dix-(Z)huit, quatre-vingt-dix (Z)ans

• **Adjectifs cardinaux « spéciaux »**

On doit toujours prononcer le **T** de ving**T** à partir de 21.

21 ans, 22 ans, 23 ans, 24 ans, 25 ans, 26 ans, 27 ans, 28 ans, 29 ans

Vin té un nan, vintdeu zan, vinttroi zan, vintqua tran, vintcin quan,

Vint si zan, vintsè tan, vintui tan, vintneu van.

9 (V)heures, 29 (V)heures, 19 (V)ans, 29 (V)ans

Cent (T)ans

**4. Les liaisons qui indiquent le nombre**

Singulier (pas de liaison)	Pluriel (liaison)
Un bois / immense	Des bois (Z)immenses
Un nez / affreux	Des nez (Z)affreux
Un prix / élevé	Des prix (Z)élevés

**5. Expressions figées**

- |                           |                     |                |
|---------------------------|---------------------|----------------|
| 1. les États-(Z)Unis      | the United States   |                |
| 2. de mieux (Z)en mieux   | better and better   |                |
| 3. de plus (Z)en plus     | more and more       | liaison avec Z |
| 4. de moins (Z)en moins   | less and less       |                |
| 5. de temps (Z)en temps   | from time to time   |                |
| 6. un fait (T)accompli    | a « fait accompli » |                |
| 7. Comment (T)allez-vous? | How are you ?       |                |
| 8. tout (T)à coup         | all of a sudden     |                |
| 9. tout (T)à fait         | entirely            | liaison avec T |
| 10. petit (T)à petit      | little by little    |                |
| 11. accent (T)aigu        | acute accent        |                |

## EXERCICES

### EXERCICE 1. Lisez les phrases suivantes

1. Il travaille de mieux en mieux.
2. Il ne faut pas d'accent aigu avant deux consonnes.
3. Petit à petit, l'oiseau fait son nid.
4. Il habite aux États-Unis.
5. Elle a de moins en moins de temps.
6. Comment allez-vous? Très bien, merci.
7. Il va voir sa grand-mère de temps en temps.
8. Il est de plus en plus beau.
9. Il pleuvait et, tout à coup, il a fait beau.
10. Cela est tout à fait réaliste.

### EXERCICE 2. Trouvez les expressions figées manquantes.

1. New-York se trouve aux \_\_\_\_\_.
2. Depuis qu'elle a suivi ce traitement, elle va de m \_\_\_\_\_.
3. Chaque jour, il est de p \_\_\_\_\_ drôle.
4. On ne peut pas tout apprendre d'un seul coup ; il faut apprendre \_\_\_\_\_ : un peu chaque jour, progressivement.
5. Elle nous a mis devant le \_\_\_\_\_.
6. Il faisait beau et \_\_\_\_\_ de gros nuages noirs sont arrivés
7. Ce n'est pas \_\_\_\_\_ (exactement) la même chose.
8. \_\_\_\_\_ l'oiseau fait son nid.
9. Lorsque les gens vieillissent, ils dorment de m \_\_\_\_\_.
10. On ne met jamais \_\_\_\_\_ ni d'accent grave avant 2 consonnes.

### Correction ex. 2

1. New-York se trouve aux États-Unis.
2. Depuis qu'elle a suivi ce traitement, elle va de mieux en mieux.
3. Chaque jour, il est de plus en plus drôle.
4. On ne peut pas tout apprendre d'un seul coup ; il faut apprendre pas à pas : un peu chaque jour, progressivement.
5. Elle nous a mis devant le fait accompli.
6. Il faisait beau et tout à coup de gros nuages noirs sont arrivés
7. Elle repassera tout à l'heure (dans un petit moment).
8. Ce n'est pas tout à fait (exactement) la même chose.
9. Lorsque les gens vieillissent, ils dorment de moins en moins.
10. On ne met jamais d'accent aigu ni d'accent grave avant deux consonnes.

## II. LES LIAISONS INTERDITES

On ne fait PAS DE LIAISON dans les cas suivants :

### 1. Nom singulier + adjectif

Un héros / élégant, un garçon / intelligent, un accident / affreux, un prix / élevé

### 2. Déterminant (article) + semi-consonne

Les / yogourts, des / yogourts ; les / yachts, des / yachts ; des / yoyos ; les yogis ; des yétis.

#### SAUF CAS PARTICULIERS

L'oiseau, un\_oiseau, cet\_oiseau, ces\_oiseaux  
                  Z                  T                  Z

Les\_oies, des\_oies  
          Z          Z

Les\_yeux, des\_yeux, aux\_yeux  
          Z          Z          Z

### 3. H aspirés (pluriel)

Noms communs les plus fréquents commençant par un H aspiré.

#### Mots féminins

1. la haine / hate
2. la hanche / hip
3. la harpe / harp
4. la hâte / hurry
5. la hausse / raise
6. la hauteur / height
7. la hiérarchie / hierarchy
8. la Hollande / Holland
9. la Hongrie / Hungary
10. la honte / shame

#### Mots masculins

1. le hamac / hammock
2. le handicap / handicap
3. le haricot / bean
4. le hasard / luck
5. le haut / top
6. le héros / hero
7. le hibou / owl
8. le homard / lobster
9. le hors-d'oeuvre / starter
10. le hurlement / screaming, howl

### EXERCICE 3

Trouvez les déterminants qui conviennent. Parfois plusieurs articles sont possibles.  
Pour cela, il faut déterminer si le H est muet (p.1) ou aspiré (p.5)

1. \_\_\_\_\_ hôtel est cher. \_\_\_\_\_ hôtels de ce quartier sont d'habitude moins chers.
2. Au zoo, Marie aime voir \_\_\_\_\_ hippopotames
3. \_\_\_\_\_ huîtres de Bretagne sont les meilleures.
4. Je déteste \_\_\_\_\_ hôpitaux.
5. \_\_\_\_\_ habitants de cette région sont très chaleureux.
6. Il a mis \_\_\_\_\_ habits du dimanche.
7. Ma mère a mal \_\_\_\_\_ hanches.
8. Astérix et Obélix sont \_\_\_\_\_ héros d'une bande dessinée.
9. \_\_\_\_\_ Hollandaises prennent \_\_\_\_\_ hors-d'œuvre.
10. \_\_\_\_\_ handicaps des francophones c'est qu'ils sont une minorité
11. \_\_\_\_\_ Hongrois habitent en Hongrie.
12. \_\_\_\_\_ hiboux hurlent dans le hamac.

### ***CORRECTION EXERCICE 3***

1. CET \_hôtel est cher. CES \_hôtels de ce quartier sont d'habitude moins chers.
2. Au zoo, Marie aime voir LES \_hippopotames
3. LES \_huîtres de Bretagne sont les meilleures.
4. Je déteste LES \_hôpital.
5. LES \_habitants de cette région sont très chaleureux.
6. Il a mis SES/ LES \_habits du dimanche.
7. Ma mère a mal AUX / hanches.
8. Astérix et Obélix sont LES /héros d'une bande dessinée.
9. LES (DES ou CES) / Hollandaises prennent DES (UN) / hors-d'œuvre.
10. LES / handicaps des Francophones c'est qu'ils sont une minorité
11. LES / Hongrois habitent en Hongrie.
12. LES / hiboux hurlent dans le hamac.

## Annexe 7 : Le texte de l'étude (236 mots):

1. Cette Anglaise a demandé aux invités les affaires des étudiants.  
E1 E2 LO1 LO2 LO3
2. Tes assistants, mes ouvriers et leurs enfants sont tous nos amis.  
LO4 LO5 LF1 LO6 LO7
3. Un bon ami a dit qu'au moyen-âge, à un certain âge, on chantait en plein air.  
LO8 LO9 LO10 LF2 LO11
4. À mon avis, ton enfant ne fait aucun effort pour s'adapter à son école.  
LO12 LO13 LF3 LO14 LF4 LO15
5. Son premier amour l'a mise devant le fait accompli, ce qui est un léger ennui.  
LO16 LO17 LF5 LO18
6. Ils sont allés au dernier étage de l'ancien édifice, mais ne sont pas allés au premier.  
LF6 LF7 LO19 LO20 LF8 LF9
7. De nouveaux étudiants ont attendu au second étage du grand immeuble.  
LO21 LI1 LF10 LO22 LO23
8. Ses vieux écrits et ses nouvelles idées lui ont valu de belles acclamations.  
LO24 LF11 LO25 LO26
9. Un oiseau aux yeux bleus regarde les oies manger des yogourts dans les yachts.  
LO27 LO28 LO29 LI2 LI3
10. Ces beaux Allemands ont eu de folles aventures grâce à leurs faux-airs de gigolos.  
LO30 LI4 LF12 LO31 E3 LO32
11. Il n'y a pas de sous-entendu : on met un accent aigu sur le « e » d'États-Unis.  
LO33 LF13 LO34 LO35 LO36
12. Cet homme a parlé d'un hôpital où il y avait de vieux habits et de vieilles éponges.  
E4 E5 LO37 E6 E7 LO38 LF14 LO39
13. Des personnes ont écrit des histoires sur ces hôpitaux.  
E8 LF15 LO40 LO41
14. Les handicaps des Hollandais et des Hongrois font d'eux des héros.  
LI5 LI6 LF16 LI7 LI8
15. Deux amis de dix ans se sont vus à six heures devant les trois arbres.  
LO42 LO43 LF17 LO44 LO45
16. Il a vingt trois ans et elle a vingt cinq ans. A neuf heures, il aura cent ans.  
E9 LO46 LO47 LI9 LI10 E10 LO48 E11 LO49 E12 LO50
17. Un héros, un garçon intelligent, a eu cet accident affreux dans un bois immense.  
LI11 LI12 E13 LI13 LO51 LI14

## Annexe 8 : Les contextes syntaxiques du texte de l'étude:

### 37 Adjectifs + noms

tes assistants(LO4)	mes ouvriers(LO5)	leurs enfants(LO6)	nos amis(LO7)
bon ami(LO8)	moyen-âge(LO9)	certain âge(LO10)	plein air(LO11)
mon avis(LO12)	ton enfant(LO13)	aucun effort(LO14)	son école(LO15)
premier amour (LO16)	léger ennui(LO18)	dernier étage(LO19)	l'ancien édifice(LO20)
nouveaux étudiants(LO21)	second étage(LO22)	un grand immeuble(LO23)	vieux écrits(LO24)
nouvelles idées(LO25)	belles acclamations(LO26)	beaux Allemands(LO30)	folles aventures (LO31)
faux-airs(LO32)	vieilles éponges(LO39)	Deux amis(LO42)	Dix ans(LO43)
Six heures(LO44)	Trois arbres(LO45)	trois ans(LO47)	neuf heures(LO49)
cent ans(LO50)			
Cette anglaise (E1)	Cet homme(E4)	cinq ans(E11)	cet accident(E13)

### 18 articles + noms

aux invités (LO1)	les affaires(LO2)	des étudiants(LO3)	Un oiseau(LO27)
aux yeux(LO28)	les oies(LO29)	un accent(LO34)	un hôpital(LO37)
vieux habits(LO38)	des histoires(LO40)	ces hôpitaux(LO41)	
des yogourts(LI1)	les yachts(LI2)	Les handicaps(LI4)	des Hollandais(LI5)
des Hongrois(LI7)	des héros(LI8)	un héros(LI11)	

### 5 Sujet (autre que pronom) + verbe

Anglaise a(E2)	homme a(E5)	personnes ont(E8)	
Allemands ont(LI4)	étudiants ont(LI1)		

### 5 noms + conjonctions (et)

ouvriers et(LF1)	Hollandais et(LF16)	ans et(LI9)	écrits et (LF11)
habits et(LF14)			

### 5 verbes + prépositions

chantait en(LF2)	s'adapter à(LF4)	allés au(LF7)	ne sont pas allés au(LF9)
vus à(LF17)			

### 4 Expressions figées

le fait accompli(LO17)	sous-entendu(LO33)	accent aigu(LO35)	États-Unis(LO36)
------------------------	--------------------	-------------------	------------------

### 4. Auxiliaire + participe passé

sont allés(LF6)	ont attendu(LF10)	ont eu(LF12)	ont écrit(LF15)
-----------------	-------------------	--------------	-----------------

### 3 verbes + articles (COD)

fait aucun(LF3)	est un(LF5)	met un(LF13)	
-----------------	-------------	--------------	--

### 3 Pronom sujet + verbe

Il a(E9)	elle a(E10)	il aura(E12)	
----------	-------------	--------------	--

### 3 Nom singulier + adjectif

un garçon intelligent(LI12)	un accident affreux(LI13)	bois immense(LI14)	
-----------------------------	---------------------------	--------------------	--

### 2 Chiffres « exceptions »

vingt trois(LO46)	vingt cinq(LO48)		
-------------------	------------------	--	--

### 1 Préposition + article

dans un (LO51)			
----------------	--	--	--

### 1 Locution prépositive

grâce à(E3)			
-------------	--	--	--

### 1 nom + adv

hôpital où(E6)			
----------------	--	--	--

### 1 pronom + pronom

il y(E7)			
----------	--	--	--

### 1 conjonction (et) + pronom

et elle (LI10)			
----------------	--	--	--

### 1 négation + participe passé

pas allés(LF8)			
----------------	--	--	--

## 51 LIAISONS OBLIGATOIRES

Les contextes réunissent à la fois les facteurs syntaxiques et phonétiques :

### 1. Articles

aux invités (LO1)                      les affaires(LO2)                      des étudiants(LO3)                      un accent(LO34)

### 2. Adjectifs possessifs

tes assistants(LO4)                      mes ouvriers(LO5)                      leurs enfants(LO6)                      nos amis(LO7)

### 3. Adjectifs avec nasales (dénasalisées)

bon ami(LO8)                      moyen-âge(LO9)                      certain âge(LO10)                      plein air(LO11)

### 4. Adjectifs avec nasales qui gardent (ou peuvent garder) la nasalité

mon avis(LO12)                      ton enfant(LO13)                      aucun effort(LO14)                      son école(LO15)

### 5. Adjectifs avec consonne latente se prononçant comme au féminin

premier amour(LO16)                      léger ennui(LO18)                      dernier étage(LO19)                      l'ancien édifice(LO20)

### 6. Adjectifs dont la consonne est dévoisée (d devient t)

second étage(LO22)                      un grand immeuble(LO23)

### 7. Adjectifs pluriels masculin (se terminant en ux)

nouveaux étudiants(LO21)                      vieux écrits(LO24)                      beaux Allemands(LO30) faux-airs(LO32)

### 8. Adjectifs pluriels féminin (se terminant en lles)

nouvelles idées(LO25)                      belles acclamations(LO26) folles aventures(LO31) vieilles éponges(LO39)

### 9. Articles avec semi- consonnes

Un oiseau(LO27)                      aux yeux(LO28)                      les oies(LO29)

### 10. Expressions figées comportant un nom

le fait accompli(LO17)                      sous-entendu(LO33)                      accent aigu(LO35)                      États-Unis(LO36)

### 11. articles ou adjectifs avec h muets

un hôpital(LO37)                      ces hôpitaux(LO41)                      vieux habits(LO38)                      des histoires(LO40)

### 12. Chiffres « réguliers » + nom

Deux amis(LO42)                      Dix ans(LO43) Six heures(LO44) Trois arbres(LO45) trois ans(LO47)

### 13. Chiffres « exceptions »

vingt trois(LO46)                      vingt cinq(LO48)                      neuf heures(LO49)                      cent ans(LO50)

### 14. Préposition + article

dans un (LO51)

## 14 LIAISONS INTERDITES

### 1. nom pluriel + verbe

étudiants ont(LI1)

### 2. déterminant + semi-consonnes

des yogourts(LI2)                      les yachts(LI3)

### 3. sujet (autre que pronom) + verbe

Allemands ont(LI4)

### 4. Déterminants + h aspirés (pluriel)

Les handicaps(LI5)                      des Hollandais(LI6)                      des Hongrois(LI7)                      des héros(LI8)

### 5. noms + et

ans et(LI9)

### 6. et + pronom

et elle(LI10)



**7. article + nom singulier avec un h aspiré** un héros(LI11)

**8. Nom singulier + adjectif**

un garçon intelligent(LI12)

un accident affreux(LI13) bois immense(LI14)

### 13 ENCHAÎNEMENTS

**1. Adjectif + nom**

Cette anglaise (E1)

Cet homme(E4)

cinq ans(E11)

cet accident(E13)

**2. Pronom sujet + verbe**

homme a(E5)

Il a(E9)

elle a(E10)

il aura(E12)

**3. Sujet (autre que pronom) + verbe**

Anglaise a(E2)

personnes ont(E8)

**4. Locution prépositive**

grâce à(E3)

**5. nom + adv**

hôpital où(E6)

**6. pronom + pronom**

il y(E7)

### 17 LIAISONS FACULTATIVES

**1. nom + et**

ouvriers et(LF1)

écrits et (LF11)

habits et (LF14)

Hollandais et(LF16)

**2. verbes + prépositions**

chantait en(LF2)

vus à(LF17)

s'adapter à(LF4)

sont allés au(LF7)

ne sont pas allés au(LF9)

**3. verbe + article (COD)**

fait aucun(LF3)

est un(LF5)

met un(LF14)

**4. Auxiliaires + participes passés**

sont allés (LF6)

ont attendu(LF11)

ont eu(LF12)

ont écrit(LF15)

**5. négation + participe passé**

pas allés(LF8)

## Annexe 9 : Fréquence des mots du texte avec *Lexique*

### mot1 fréquence moins de 10

LF12	écrits	2	a
LI6	Hollandais	2	a
LF4	adapter	5	a
LO31	folles	6	a
LO36	Etats	8	a
LF2	chantait	8	a

### mot1 fréquence 11-20

LO35	accent	12	B
LF1	ouvriers	14	B
LO22	second	15	B
LO18	léger	17	B
LO39	vieilles	17	B
LF15	habits	19	B

### mot1 fréquence 21-40

LF10	étudiants	21	C
LF7	allés	26	C
LF9	allés	26	C
LO46	vingt	29	C
LO48	vingt	29	C
LI3	Allemands	30	C
LO20	ancien	32	C
LO49	neuf	32	C
LO25	nouvelles	38	C
LO21	nouveaux	39	C

### mot1 fréquence 41-60

LO9	moyen	41	D
LO50	cent	43	D
LF17	vus	47	D
LO26	belles	49	D
E2	Anglaise	51	D
LO13	ton	51	D
LO30	beaux	57	D

### mot1 fréquence 61-80

LO10	certain	65	E
------	---------	----	---

### mot1 fréquence 81-100

E3	grâce	85	F
LF14	met	89	F
LO32	faux	90	F
LI13	accident	100	F

### mot1 fréquence 101-200

LO11	plein	104	G
------	-------	-----	---

E8	personnes	105	G
LI14	bois	115	G
LO44	six	117	G
LO43	dix	118	G
E6	hôpital	126	G
LO19	dernier	138	G
LO16	premier	146	G
E11	cinq	161	G
LO14	aucun	175	G
LO24	vieux	180	G
LO38	vieux	180	G
LI12	garçon	188	G

mot1 fréquence 201-500

LO6	leurs	243	H
LO33	sous	315	H
LO23	grand	338	H
LO45	trois	380	H
LO47	trois	380	H
LO17	fait	384	H
E4	cet	499	H
E13	cet	499	H

mot1 fréquence 501-1000

LF6	sont	511	I
LO7	nos	524	I
LO4	Tes	690	I
LI9	ans	718	I
E5	homme	781	I
LO1	aux	835	I
LO28	aux	835	I
LO8	bon	836	I

mot1 fréquence 1001-5000

LO42	Deux	1009	J
LF11	ont	1063	J
LF13	ont	1063	J
LF16	ont	1063	J
LO5	mes	1102	J
LO41	ces	1222	J
LO15	son	1740	J
E1	cette	1902	J
LF3	fait	2751	J
LO12	mon	3848	J
E10	elle	4520	J
LO51	dans	4658	J

mot1 fréquence 5001-6500

LO3	des	6055	K
-----	-----	------	---

LO40	des	6055	K
LI1	des	6055	K
LI5	des	6055	K
LI7	des	6055	K
LI8	des	6055	K

mot1 fréquence 6501 à 8000

0 ----- 1

mot1 fréquence plus de 8001

LO2	les	8720	M
LO29	les	8720	M
LI2	les	8720	M
LI4	les	8720	M
LO27	un	12087	M
LO34	un	12087	M
LO37	un	12087	M
LI11	un	12087	M
LI10	et	12909	M
E7	il	13222	M
E9	il	13222	M
E12	il	13222	M
LF8	pas	18188	M
LF5	est	19417	M

**Annexe 10 : Tableau fréquence du mot1 et du mot2 dans *Vocabprofil* - fréquences cumulées.**

<b>tableau 1</b>	<b>enchaînements</b>	<b>VOC</b>	<b>VOC</b>	<b>ind.</b>	<b>ind.</b>	<b>Total</b>
		<b>M1</b>	<b>M2</b>	<b>M1</b>	<b>M2</b>	
E1	cette Anglaise	K1	K1	3	3	6
E2	Anglaise a	K1	K1	3	3	6
E3	grâce à	K1	K1	3	3	6
E4	cet homme	K1	K1	3	3	6
E5	homme a	K1	K1	3	3	6
E6	hôpital où	K1	K1	3	3	6
E7	il y	K1	K1	3	3	6
E8	personnes ont	K1	K1	3	3	6
E9	il a	K1	K1	3	3	6
E10	elle a	K1	K1	3	3	6
E11	cinq ans	K1	K1	3	3	6
E12	il aura	K1	K1	3	3	6
E13	cet accident	K1	K1	3	3	6
<b>tableau 2</b>	<b>Liaisons obligatoires</b>	<b>VOC</b>	<b>VOC</b>	<b>ind.</b>	<b>ind.</b>	<b>Total</b>
		<b>M1</b>	<b>M2</b>	<b>M1</b>	<b>M2</b>	
LO8	bon ami	K1	K1	3	3	6
LO7	nos amis	K1	K1	3	3	6
LO6	leurs enfants	K1	K1	3	3	6
LO51	dans un	K1	K1	3	3	6
LO50	cent ans	K1	K1	3	3	6
LO49	neuf heures	K1	K1	3	3	6
LO48	vingt cinq	K1	K1	3	3	6
LO47	trois ans	K1	K1	3	3	6
LO46	vingt trois	K1	K1	3	3	6
LO45	trois arbres	K1	K1	3	3	6
LO44	six heures	K1	K1	3	3	6
LO43	dix ans	K1	K1	3	3	6
LO42	Deux amis	K1	K1	3	3	6
LO41	ces hôpitaux	K1	K1	3	3	6
LO40	des histoires	K1	K1	3	3	6
LO4	Tes assistants	K1	K1	3	3	6
LO37	un hôpital	K1	K1	3	3	6
LO34	un accent	K1	K1	3	3	6
LO33	sous-entendu	K1	K1	3	3	6
LO32	faux-airs	K1	K1	3	3	6
LO30	beaux Allemands	K1	K1	3	3	6
LO3	des étudiants	K1	K1	3	3	6
LO28	aux yeux	K1	K1	3	3	6
LO27	un oiseau	K1	K1	3	3	6
LO25	nouvelles idées	K1	K1	3	3	6
LO24	vieux écrits	K1	K1	3	3	6
LO22	second étage	K1	K1	3	3	6

LO21	nouveaux étudiants	K1	K1	3	3	6
LO2	les affaires	K1	K1	3	3	6
LO19	dernier étage	K1	K1	3	3	6
LO16	premier amour	K1	K1	3	3	6
LO15	son école	K1	K1	3	3	6
LO14	aucun effort	K1	K1	3	3	6
LO13	ton enfant	K1	K1	3	3	6
LO12	mon avis	K1	K1	3	3	6
LO11	plein air	K1	K1	3	3	6
LO1	aux invités	K1	K1	3	3	6
LO9	moyen-âge	K1	K2	3	2	5
LO5	mes ouvriers	K1	K2	3	2	5
LO36	Etats-Unis	K1	K2	3	2	5
LO23	grand immeuble	K1	K2	3	2	5
LO10	certain âge	K1	K2	3	2	5
LO31	folles aventures	K2	K2	2	2	4
LO20	ancien édifice	K1	K3	3	1	4
LO17	fait accompli	K1	K3	3	1	4
LO39	vieilles éponges	K1	HL	3	0	3
LO38	vieux habits	K1	HL	3	0	3
LO29	les oies	K1	HL	3	0	3
LO26	belles acclamations	K1	HL	3	0	3
LO18	léger ennui	K1	HL	3	0	3
LO35	accent aigu	K3	HL	1	0	1

<b>tableau 3</b>	<b>Liaisons facultatives</b>	<b>VOC</b>	<b>VOC</b>	<b>ind.</b>	<b>ind.</b>	<b>Total</b>
LF3	fait aucun	K1	K1	3	3	6
LF5	est un	K1	K1	3	3	6
LF6	sont allés	K1	K1	3	3	6
LF7	allés au	K1	K1	3	3	6
LF8	pas allés	K1	K1	3	3	6
LF9	allés au	K1	K1	3	3	6
LF10	étudiants ont	K1	K1	3	3	6
LF11	ont attendu	K1	K1	3	3	6
LF12	écrits et	K1	K1	3	3	6
LF13	ont eu	K1	K1	3	3	6
LF14	met un	K1	K1	3	3	6
LF16	ont écrit	K1	K1	3	3	6
LF17	vus à	K1	K1	3	3	6
LF1	ouvriers et	K2	K1	2	3	5
LF2	chantait en	K2	K1	2	3	5
LF4	s'adapter à	K2	K1	2	3	5
LF15	habits et	HL	K1	0	3	3

<b>tableau 4</b>	<b>liaisons interdites</b>	<b>VOC</b>	<b>VOC</b>	<b>ind.</b>	<b>ind.</b>	<b>Total</b>
LI3	Allemands ont	K1	K1	3	3	6
LI9	ans et	K1	K1	3	3	6
LI10	et elle	K1	K1	3	3	6

LI8	des héros	K1	K2	3	2	<b>5</b>
L I11	Un héros	K1	K2	3	2	<b>5</b>
LI12	garçon intelligent	K1	K2	3	2	<b>5</b>
LI14	bois immense	K1	K2	3	2	<b>5</b>
LI1	des yogourts	K1	HL	3	0	<b>3</b>
LI2	les yachts	K1	HL	3	0	<b>3</b>
LI4	Les handicaps	K1	HL	3	0	<b>3</b>
LI5	des Hollandais	K1	HL	3	0	<b>3</b>
LI6	Hollandais et	K1	HL	3	0	<b>3</b>
LI7	des Hongrois	K1	HL	3	0	<b>3</b>
L I13	accident affreux	K1	HL	3	0	<b>3</b>

## Annexe 11 : Exemples de codage *GoldVarb* pour fréquence avec *Lexique* et *Vocabprofil*

**c** : code donné à toutes les occurrences (chaque occurrence a un code qui lui est unique)

**sand** : type de sandhi. e=enchaînement, o =liaison obligatoire, f=liaison facultative, i = liaison interdite

**Ph** : contexte phonétique

**synt** : contexte syntaxique

**% L** : Type de sandhi (plus détaillé : ex : 4 = avec nasale)

**Lex. M1** = fréquence du mot1 avec *Lexique*

**voc. M1** = fréquence du mot1 avec *Vocabprofil*

**Lex. M2** = fréquence du mot2 avec *Lexique*

**voc. M2** = fréquence du mot1 avec *Vocabprofil*

**Lex. cooc** = fréquence de cooccurrence avec *Lexique*

	c oc	sand	Ph.	% L	synt	Lex.M1	voc.M1	Lex.M2	voc.M2	Lex.
cette Angl.	a	e	t	/	D	j	3	d	3	A
Anglaise a	à	e	z	/	S	d	3	k	3	A
aux invités	â	o	z	5	P	i	3	c	3	C
les affaires	b	o	z	5	A	m	3	g	3	M
des étud.	c	o	z	5	A	k	3	c	3	H
Tes assist.	ç	o	z	5	P	i	3	a	3	A
mes ouv.	d	o	z	5	P	j	3	b	2	A
ouvriers et	e	f	z	5	E	b	3	m	3	B
leurs enf.	é	o	z	5	P	h	3	h	3	I
nos amis	è	o	z	5	P	i	3	h	3	K
bon ami	ê	o	n	4	Q	i	3	h	3	I
moyen-âge	f	o	n	4	Q	d	3	g	2	C
certain âge	g	o	n	4	I	e	3	g	2	D
chantait en	h	f	t	3	U	a	2	k	3	A
plein air	i	o	n	4	Q	g	3	h	3	G
mon avis	î	o	n	4	P	j	3	g	3	M
ton enfant	ï	o	n	4	P	d	3	h	3	G
fait aucun	j	f	t	3	V	j	3	g	3	H
aucun effort	k	o	n	4	I	g	3	c	3	C
s'adapter à	l	f	r	2	U	a	2	m	3	E
son école	m	o	n	4	P	j	3	g	3	E
premier am.	n	o	r	2	O	g	3	h	3	F
fait acc.	o	o	t	3	N	h	3	a	1	B
est un	ô	f	t	3	T	m	3	m	3	M
léger ennui	p	o	r	2	Q	b	2	b	0	A
sont allés	q	f	t	3	X	i	3	c	3	I
allés au	r	f	z	5	U	c	3	j	3	G
dernier ét.	s	o	r	2	Q	g	3	c	3	F
ancien éd.	t	o	n	4	Q	c	3	a	1	A
pas allés	u	f	z	5	G	m	3	c	3	F
allés au	ù	f	z	5	U	c	3	j	3	G
Nouv. étud.	û	o	z	5	Q	c	3	c	3	B
Étud.ont	v	i	z	5	S	c	3	j	3	C
ont attendu	w	f	t	3	X	j	3	c	3	D
Sec.étage	x	o	t	3	O	b	3	c	3	B
grand imm.	y	o	t	3	Q	h	3	c	2	B
vieux écrits	z	o	z	5	Q	g	3	a	3	A
écrits et	A	f	z	5	E	a	3	m	3	B
Nouv.idées	À	o	z	5	Q	c	3	d	3	C
belles accl.	Â	o	z	5	Q	d	3	a	0	A
un oiseau	B	o	n	4	A	m	3	d	3	K
aux yeux	C	o	z	5	P	i	3	h	3	L
les oies	Ç	o	z	5	A	m	3	a	0	C



des yog.	D	i	0	5	A	k	3	a	0	A
les yachts	E	i	0	5	A	m	3	a	0	A
beaux All.	É	o	z	5	Q	d	3	c	3	A
Allem. ont	È	i	0	5	S	c	3	j	3	E
ont eu	Ê	f	t	3	X	j	3	i	3	M
folles av.	F	o	z	5	Q	a	2	a	2	A
grâce à	G	e	s	/	L	f	3	m	3	M
faux-airs	H	o	z	5	Q	f	3	b	3	A
sous-ent.	I	o	z	5	Z	h	3	h	3	C
met un	Î	f	t	3	V	f	3	m	3	G
un accent	Ï	o	n	4	A	m	3	b	3	G
accent aigu	J	o	t	3	N	b	1	a	0	A
Etats-Unis	K	o	z	5	N	a	3	a	2	M
cet homme	L	e	t	/	D	h	3	i	3	M
homme a	M	e	m	/	S	i	3	k	3	M
un hôpital	N	o	n	4	A	m	3	g	3	K
hopital où	O	e	l	/	U	g	3	j	3	D
il y	Ô	e	l	/	Y	m	3	j	3	M
vieux habits	P	o	z	5	Q	g	3	b	0	B
habits et	Q	f	z	5	E	b	0	m	3	D
vieilles ép.	R	o	z	5	Q	b	3	a	0	A
pers.ont	S	e	n	/	S	g	3	j	3	H
ont écrit	T	f	t	3	X	j	2	g	3	F
des histoires	U	o	z	5	A	k	3	e	3	M
ces hôpitaux	Û	o	z	5	D	j	3	a	3	A
Les hand.	V	i	0	5	A	m	3	a	0	A
des Holl.	W	i	0	5	A	k	3	a	0	A
Holl.et	X	i	0	5	E	a	0	m	3	A
des Hongr.	Y	i	0	5	A	k	3	a	0	A
des héros	Z	i	0	5	A	k	3	e	2	H
Deux amis	0	o	z	5	N	j	3	h	3	G
dix ans	1	o	z	5	N	g	3	i	3	M
vus à	2	f	z	5	U	d	3	m	3	E
six heures	3	o	z	5	N	g	3	h	3	I
trois arbres	4	o	z	5	N	h	3	c	3	A
il a	5	e	l	/	L	m	3	k	3	M
vingt trois	6	o	t	3	C	c	3	h	3	E
trois ans	7	o	z	5	N	h	3	i	3	M
ans et	8	f	0	5	E	i	3	m	3	M
et elle	9	i	0	3	E	m	3	j	3	M
elle a	et	e	l	/	L	j	3	k	3	M
vingt cinq	(	o	t	3	C	c	3	g	3	G
cinq ans	)	e	k	/	N	g	3	i	3	M
neuf heures	\$	o	v	0	N	c	3	h	3	G
il aura	£	e	l	/	L	m	3	g	3	M
cent ans	μ	o	t	3	N	d	3	i	3	H
Un héros	[	i	0	4	A	m	3	e	2	M
garçon int.	]	i	0	4	N	g	3	c	2	C
cet accident	/	e	t	/	D	h	3	f	3	G
accident aff.	*	i	0	3	N	f	3	c	0	A
dans un	!	o	z	5	P	j	3	m	3	M
bois imm.	?	i	0	5	N	g	3	b	2	A

## Bibliographie

- Abe, H. (2006). Input Enhancement in Pronunciation Pedagogy: The Impact on Learning Connected Speech in L2 English. *Research Reports of Tsuruoka National College of Technology*, 41, 15-38.
- Abry, D., et Chalaron, M.L. (1994). *Phonétique – 350 exercices- hachette F.L.E.*
- Acton, W. (1997). Seven Suggestions of Highly Successful Pronunciation Teaching. The Language Teacher Online. <http://jalt-publications.org/tlt/files/97/feb/seven.html>
- Adda-Decker, M., Boula de Mareüil, P., et Lamel, L. (1999). Pronunciation variants in French: schwa & liaison, *ICPhS*, 2239-2242, San Francisco.
- Ågren, J. (1973). *Étude sur quelques liaisons facultatives dans le français de conversation radiophonique: fréquences et facteurs*. Uppsala: Uppsala University Press.
- Ahmad, M. (1993). Vingt heures de français parlé. Aspects phonétiques de la liaison. (*Thèse de doctorat, Université Grenoble 3*). Grenoble.
- Akahane-Yamada, R., Tohkura, Y., Bradlow, A., et Pisoni, D. (1996). Does training in speech perception modify speech production? *4<sup>th</sup> International Conference on Spoken Language Processing-1996*, 606-609.  
<http://www.asel.udel.edu/icslp/cdrom/vol2/277/a277.pdf>
- Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, 26, 339–356.
- Armstrong, N. (2001). Social et stylistic variation in spoken French: a comparative approach. *Chapter 5. Variable liaison*. (pp.177-207). Amsterdam: John Benjamins Pub. Co.
- Ashby, M., Maidment, J., et Abberton, E. (1996). Analytic listening: a new approach to ear-training, 5-8.
- Ashby, W. J. (1981). French liaison as a sociolinguistic phenomenon. W. Cressey and D.J. Napoli (Eds.), *Linguistic Symposium on Romance Languages: 9*. Washington, D.C.: Georgetown University Press, pp. 46-57.
- Ashby W. J. (1988). Français du Canada/français de France: divergence et convergence. *The French Review*, 61 (5), 693-702.
- Ashby, W. J. (2003). La liaison variable en français parlé tourangeau : une analyse en temps réel *Colloque AFLS Tours : 25-27 septembre 2003*.
- Azra, J.-L. (2004). Découpe syllabique, liaisons, enchaînements, rythme. - Quelle progression de classe pour la prononciation et la compréhension orale ? *Rencontres Pédagogiques du Kansai. Thème 1*. [http://www.rpkansai.com/bulletins/pdf/018/008\\_010\\_azura.pdf](http://www.rpkansai.com/bulletins/pdf/018/008_010_azura.pdf)
- Bardovi-Harlig, K. (1987). Markedness and salience in second language acquisition. *Language Learning* 37, 385-407.

- Bartning, I. (1997) : L'apprenant dit avancé et son acquisition d'une langue étrangère. Tour d'horizon et esquisse de la variété avancée. In Bartning, I. (ed), *Acquisition et Interaction en Langue Etrangère (AILE) 9 (n° spécial, Les Apprenants Avancés)*, 9-50.
- Bastin, E. (2007). La méthode Balingua ou comment démystifier l'apprentissage des langues. *Science et Culture*. Bulletin n° 407 de Mai-Juin 2007, 1-10.  
<http://www.universalapproach.com/library/Documents/EB%20ULg%20avril%202007.pdf>
- Bates, E., et MacWhinney, B. (1979). A functionalist approach to the acquisition of grammar. In E. Ochs and B. Schieffelin (Eds.), *Developmental Pragmatics*. Academic Press, Inc., New York.
- Bates, E., et MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*. New York: Cambridge University Press.
- Bates, E. et MacWhinney, B. (1987). Competition, variation, and language learning. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition* (pp.157-193). Mahwah, NJ: Erlbaum.
- Bayley, R. et Preston, D.R. (1996). *Second Language Acquisition and Linguistic Variation*, John Benjamins.
- Bayley, R. (1994). Interlanguage variation and the quantitative paradigm: Past tense marking in Chinese English. In E. Tarone, S. M. Gass, & A. D. Cohen (Eds.), *Research Methodology in Second Language Acquisition* (pp.157-181). Mahwah, NJ: Erlbaum.
- Beaudoin, M. (1998). Découpage syllabique en français comme langue seconde. *The Canadian Modern Language Review/La Revue Canadienne des Langues Vivantes*, 54 (3), 354-375.
- Beeching, K. (2001). Un corpus d'entretiens spontanés, enregistrés et transcrits par Kate Beeching. [En ligne] Disponible sur <http://www.uwe.ac.uk/facults/les/staff/kb/CORPUS.pdf>
- Belletti, A. et Hamann, C. (2000). Ça on fait pas! On the L2-acquisition of French by two young children with different source languages. In S.C. Howell, S.A. Fish & T. Keith-Lucas (Eds.), *Proceedings of the 24th annual Boston University conference on language development*. Somerville, MA: Cascadilla Press, pp. 116-127.
- Belletti, A. et Hamann, C. (2004). On the L2/bilingual acquisition of French by two young children with different source languages. In P. Prévost & J. Paradis, (Eds.), *The Acquisition of French in Different Contexts: Focus on Functional Categories*. Amsterdam: John Benjamins.
- Berri, A. (2007). Quelques considérations à propos de la phonétique dans la formation des futurs enseignants de FLE. *Fragmentos, número 33*, p. 245/250 Florianópolis/ jul - dez/ 2007  
<http://www.periodicos.ufsc.br/index.php/fragmentos/article/viewFile/8667/8008>
- Bild, E.R. et Swain, M. (1989). Minority language students in a French immersion programme: their French proficiency. In: *Journal of Multilingual and Multicultural Development* 10 (3), 255-274.

- Birdsong, D. (2003). Authenticity in French as a second language pronunciation by adult English speaking learners: Segmental and global analyses. *Aile: Acquisition et Interaction en Langue Étrangère*, 18, 17-36.
- Blanche, P. (2004). *Linguagem & Ensino*, 7 (1), 175-191. Using dictation to teach pronunciation.  
[http://rle.ucpel.tche.br/php/edicoes/v7n1/I\\_patrick2.pdf](http://rle.ucpel.tche.br/php/edicoes/v7n1/I_patrick2.pdf)
- Boë, L.-J. et Tubach, J.-P. (1992). *De A à Zut: dictionnaire phonétique du français parlé*. Grenoble: Ellug.
- Bongaerts, T. (2003). The effects of Age in the Acquisition of pronunciation of a second language. *Aile: Acquisition et Interaction en Langue Étrangère*, 18, 79-98.
- Booij, G., et Jong, D. (1987). The domain of liaison: theories and data. *Linguistics*, 25, 1005–1025.
- Boubnova, G. (2000). Contrastive prosody of Russian and French: models and applications. *Research supports scheme electronic library*.
- Boula de Mareüil, P., et Adda-Decker, M. (2002). Studying pronunciation variants in French by using alignment techniques. *ICSLP*, Denver. In *ICSLP-2002*, 2273-2276.  
<http://www.limsi.fr/Individu/mareuil/publi/SL021383.pdf>
- Boula de Mareüil, P., Adda-Decker, M., et Gendner, V. (2003). Liaisons in French: a corpus-based study using morpho-syntactic information. 15<sup>th</sup> ICPHS, 1329-1332, Barcelona.  
<http://www.limsi.fr/Individu/mareuil/publi/PS020851.pdf>
- Bourdages, J. S. et Champagne-Muzar, C. (1993). *Le point sur la phonétique*. Paris : Clé International, 1998.
- Broner, M. (2001). *Impact of interlocutor and task on first and second language use in a Spanish Immersion program* (CARLA, Working Paper No. 18, Summarized in:  
<http://www.carla.umn.edu/immersion/broner.html>). Minneapolis, MN: CARLA.
- Brooks, P. et Tomasello, M. (1999). Young children learn to produce passives with nonce verbs. *Developmental Psychology*, 35, 29-44.
- Bybee, J. (1985). *Morphology*. Amsterdam: John Benjamins.
- Bybee, J. (1995). The development of past tense modals in English. In J. Bybee and S. Fleischman (Eds.), *Modality in Grammar and Discourse* (pp. 503-17). Amsterdam: John Benjamins.
- Bybee, J. (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication*, 2 (2), 249-278.
- Bybee, J. (2001). Frequency effects on French liaison. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 337-359). Amsterdam: Benjamins.

- Bybee, J. (2002). Phonological evidence for exemplar storage of multiword sequences. *Studies in Second Language Acquisition*, 24 (2), 215-221
- Bybee, J. (2005). Liaison: effets de fréquence et constructions. *Langages*, 158, 24-37.
- Bybee, J., et Scheibman, J. (1999). The effect of usage on degree of constituency: The reduction of don't in American English. *Linguistics*, 37, 575-596.
- Bybee, J. et Thompson, S. (1997). Three frequency effects in syntax. *Berkeley Linguistic Society*, 23, 378-388.
- Bybee, J. and S. Thompson, S. (2000). Three frequency effects in syntax. *Berkeley Linguistic Society*, 23, 65-85.
- Cappeau, P., et Gadet, F. (2007). Document 3. Où en sont les corpus sur les français parlés ? *Revue Française de Linguistique Appliquée* 2007/1, 121, 129-133.
- Catford, J.C. et Pisoni, D.B. (1970). Auditory vs articulatory training in exotic sounds. *Modern Language Journal*. 54 (7), 447-481.
- Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, 7 (1), 71-87.
- Cenoz, J. et Lecumberri, L.G. (1999). The effect of training on the discrimination of English vowels. *International Review of Applied Linguistics* 37 (4), 261-275.
- Chabanal, D. (2003). *Un aspect de l'acquisition du français oral : la variation socio-phonétique chez l'enfant francophone*. Doctorat de Sciences du langage, Université Paul Valéry, Montpellier III.
- Chabanal, D. (2007). L'apport de la sociolinguistique développementale pour l'apprentissage de l'oral. Dixième colloque international de l'Association internationale pour la recherche en didactique du français (AIRDF). *Didactique du français : le socioculturel en question*. 13-15 septembre 2007
- Chabanal, D. et Embarki, M. (2002). L'acquisition d'un marqueur socio-stylistique : L'exemple de la liaison facultative. XXIVèmes Journées d'Étude sur la Parole, Nancy, 24-27 juin 2002
- Charliac, L., et Motron, A.-C. (1998). *Phonétique progressive du français avec 600 exercices*. [France] : CLE International.
- Chevrot, J.P. (2002). Acquisitions langagières chez l'enfant d'école maternelle : le cas de la segmentation des mots et des conduites multimodales. *Projet scientifique de l'équipe Acquisitions Langagières et Usages Linguistiques Infantins, LIDILEM*, Grenoble III.
- Chevrot, J.P. et Fayol, M. (2000). L'acquisition de la liaison: enjeux théoriques, premiers résultats, perspectives. *Lidil*, 22, 11-30.

- Chevrot, J.-P., Dugua, C., et Fayol, M. (2005). Liaison et formation des mots en français; un scénario développemental. *Langages*, 158, 38-52.
- Chevrot, J.-P., Chabanal, D., et Dugua, C. (2007). Pour un modèle de l'acquisition des liaisons basé sur l'usage: trois études de cas. *Journal of French Language Studies*, 17 (1), 103-128.
- Chevrot, J.-P., Fayol, M., et Laks, B. (2005). La liaison: de la phonologie à la cognition. *Langages*, 158, 3-7.
- Choi, J.-N. (2003). A Typological Comparison of Complex Korean and French Statements: Their development in oral narratives of French-Korean children in consecutive bilingual acquisition. *Aile: Acquisition et interaction en langue étrangère*, 2003, 19, Apr (2003), 107-46.
- Chomsky, N. (1959). Review of B. F. Skinner's Verbal Behavior. *Language*, 35 (1), 26-58.
- Clements, G. N. et Keyser, S. J. (1981). A three-tired theory of the Syllable. In C. F.C. Science (Ed.), *Occasional Papers 19*. Cambridge : MA : MIT Press.
- Clements, G. N. et Keyser, S. J. (1983). *CV Phonology, A Generative Theory of Syllable*. Cambridge Massachusetts : MIT Press.
- Couper, G. (2003). The value of an explicit pronunciation syllabus. *ESOL Teaching*. Prospect 18 (3), 53-70
- Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect* 21 (1), 46-66
- Culetto, T. (2007). Prediction of liaison in French by measures of Information Theory. *LingO* 2007.
- Curtiss, S. (1977). *Genie: A Psycholinguistic Study of a Modern Day : "Wild Child"*. New York Academic Press.
- Cutler, A., Mehler, J., Norris, D. et Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25, 385-400.
- Cutler, A., Mehler, J., Norris, D., et Seguí, J. (1989). Limits on bilingualism. *Nature*, 340, 229-230.
- Dansereau, D. (1995). Phonetics in the beginning and intermediate oral proficiency-oriented French classroom. *The French Review*, 68 (4), 638-651.
- De Jong, D. (1988). *Sociolinguistic aspects of French Liaison*. Dordrecht : Foris.
- De Jong, D. (1989). A Multivariate analysis of French liaison, in M.E.H.S. Schouten and Th. Van Reenen-Pieter (Eds.), *New Methods in Dialectology*, Dordrecht: Foris, 19-32.

- De Jong, D. (1991). La liaison à Orléans et à Montréal, *Proceedings of the XIIIth international Congress of Phonétic Sciences*, 315, 198-201.
- De Jong, D. (1994). La sociophonologie de la liaison orléanaise. In: C. Lyche (Ed.), *French Generative Phonology : Restrospective and Perspectives*, AFLS/ESRI, 95-130.
- Delattre, P. (1946). Pour imiter un disque de français parlé . *The French Review*, 20 (1), 43-48.
- Delattre, P. (1947). La liaison en français, tendances et classification. *The French Review*, 21 (2), 148-157.
- Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants anglo-américains*. Vermont : Middlebury College, 2<sup>e</sup> édition.
- Delattre, P. (1955). Les facteurs de la liaison facultative en français. *The French Review*, 29 (1), 42-49.
- Delattre, P. (1956). La fréquence des liaisons facultatives en français. *The French Review*, 30 (1), 48-54.
- Delattre, P. 1965. *Comparing the Phonetic Features of English, French, German and Spanish: An Interim Report*. Heidelberg: Julius Groos Verlag.
- Delattre, P. (1966). *Studies in French and Comparative Linguistics*, The Hague, Mouton.
- Delattre, P. (1969). Syllabic features and phonic impression in English, German, French and Spanish. *Lingua*, 22, 160-175.
- Dell, F. (1970). *Les règles phonologiques tardives et la morphologie dérivationnelle du français*, Doctoral Dissertation, Cambridge, MIT Press.
- Dell, F. (1973). *Les règles et les sons*. Paris : Hermann.
- Derwing, T. M., et Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency and complexity of L2 accented speech. *Applied Language Learning*, 13, 1-17.
- Desrochers, R. (1994). Les liaisons dangereuses : le statut équivoque des erreurs de liaison. *Linguisticae investigationes*, 18 (2): 243-284.
- Detey, S., et Nespoulous, J.-L. (2008). Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua*, 118 (1), 66-81.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25, 108-127.
- Domergue, P. (1805). *Manuel des étrangers amateurs de la langue française*. Paris, Librairie économique.

- Dubois, J., Giacomo, M., Gespin L., Marcellesi, C., Marcellesi, J.B., Mével, J.P., (2007). Grand dictionnaire de linguistique et des sciences du langage. Paris, Larousse
- Dufeu, S., Lété, B., Touzet, C., Glotin, H., Ziegler, J.C., et Grainger, J. (2006). Modélisation et simulation par carte auto-organisatrice de l'effet de fréquence des mots chez l'apprenti lecteur. *Proceedings of the First French Conference on Computational Neuroscience*. Pont-à-Mousson, France.  
[http://recherche.univ-lyon2.fr/emc/IMG/pdf/2006\\_Dufau\\_al\\_Neurocomp2006.pdf](http://recherche.univ-lyon2.fr/emc/IMG/pdf/2006_Dufau_al_Neurocomp2006.pdf)
- Dufeu, B. (1986). Rythme et expression. Apprentissage de l'intonation : quand le son est aussi porteur de sens, dans *Le Français dans le monde*, 205, 62-70.
- Dufeu, B. (2008). L'importance de la prononciation dans l'apprentissage d'une langue étrangère. *Franc-parler*, 1-10.  
[http://www.franparler.org/dossiers/phonetique\\_dufeu3.htm](http://www.franparler.org/dossiers/phonetique_dufeu3.htm)
- Dugua, C. (2002). *Liaison et segmentation du lexique en français : vers un scénario développemental*, Mémoire de DEA en Sciences du langage, Université Stendhal, Grenoble.
- Dugua, C. (2005). De la liaison à la formation du lexique chez les jeunes enfants francophones. *Le langage et l'homme*, 40 (2), 163-182.
- Dugua, C. (2006). "un nous"/"des nous" ou le rôle de la fréquence sur l'acquisition de la liaison en français. In Loiseau et al. (Eds.). *Autour des langues et du langage : perspective pluridisciplinaire*, Papiers sélectionnés du Colloque international des étudiants chercheurs en didactique des langues et en linguistique, Grenoble : PUG.
- Dugua, C., et Chabanal, D. (2006). Acquisition de la liaison chez l'enfant francophone : formes lexicales des Mots 2. Présentation de communication. XXVI<sup>es</sup> Journées d'Étude sur la Parole, Dinard, 12-16 juin 2006.
- Duménil, A. (2003). *Facile à dire ! Les sons du français*. Upper Saddle River, N.J.: Prentice Hall.
- Durand, M. (1936). *Le genre grammatical en français parlé à Paris et dans la région parisienne*. Paris: D'Artrey.
- Durand, J., Laks, B., et Lyche, C. (2002). La phonologie du français contemporain: usages, variétés et structure. In: C. Pusch & W. Raible (Eds.) *Romanistische Korpuslinguistik-Korpora und gesprochene Sprache/Romance Corpus Linguistics - Corpora and Spoken Language*. Tübingen: Gunter Narr Verlag, pp. 93-106.
- Durand, J., Laks, B., et Lyche, C. (2005). Un corpus numérisé pour la phonologie du français. In G. Williams (Ed.) *La linguistique de corpus*. Rennes: Presses Universitaires de Rennes. pp. 205-217. Actes du colloque 'La linguistique de corpus', Lorient, 12-14 septembre 2002.
- Durand, J., et Lyche, C. (2008). French liaison in the light of corpus data. *Journal of French Language Studies*, 18 (1), 33-66.



- Elliott, A. R. (1994). Predicting near-native pronunciation in Spanish as a foreign language. *UTA. Working Papers in Linguistics*, 1, 51-68.
- Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188.
- Ellis, N. C., et Laporte, N. (1997). Contexts of acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In A. M. B. de Groot and J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*, 53-83. Hillsdale, N. J. Lawrence Erlbaum.
- Encrevé, P. (1983). La liaison sans enchaînement. *Actes de la recherche en sciences sociales*, 46, 39-66.
- Encrevé, P. (1988). *La liaison avec et sans enchaînement. Phonologie tridimensionnelle et usages du français*. Paris : Seuil
- Enomoto, K. (1994). L2 *Perceptual Acquisition: The Effect of Multilingual Linguistic Experience on the Perception of a "Less Novel" Contrast*. Edinburgh Working Papers in *Applied Linguistics*, 5, 15-29.
- Erman, B. et Warren, B. (2000). The Idiom Principle and the Open Choice Principle. *Text*, 20, 29-62.
- Eskildsen, S. W. (2009). Constructing another language –Usage-Based linguistics in second language acquisition. *Applied Linguistics*, 30(3):335-357; *Applied Linguistics Advance* Access originally published online on November 11, 2008. Oxford University Press 2008. <http://applied.oxfordjournals.org/cgi/content/full/amn037>
- Esposito, F. (1994). *Les liaisons dangereuses : pour ne plus lier les mots à tort et à travers*. Paris, Victoire Éditions, 1994.
- Ferrand, L., Grainger, J., et New, B. (2003). Normes d'âge d'acquisition pour 400 mots monosyllabiques. *L'Année Psychologique*, 104, 445-468.
- Fouché (Eds.). (1959). *Traité de prononciation française*. Paris: Klincksieck.
- Fougeron, C., Bagou, O., Stefanuto, M. et Frauenfelder, U. H. (2002). A la recherche d'indices de frontière lexicale dans la resyllabation. *Actes des XXIV<sup>es</sup> Journée d'Etude sur la Parole*, Nancy, France, pp 125-128.
- Fougeron, C., et Delais-Roussarie, E. (2004). Liaisons et enchaînements : « Fais\_en à Fez\_en parlant ». *Paper presented at Actes des Journées d'Etudes sur la Parole 2004*, Fès, Morocco. [http://ed268.univ-paris3.fr/lpp/pages/EQUIPE/fougeron/CF\\_pub/Fougeron-Delais-JEP04.pdf](http://ed268.univ-paris3.fr/lpp/pages/EQUIPE/fougeron/CF_pub/Fougeron-Delais-JEP04.pdf)

- Fougeron, C., Goldman, J.-P., et Frauenfelder, U. (2001). Liaison and schwa deletion in French: an effect of lexical frequency and competition. *Eurospeech conference*. Aalborg, Danemark.
- Fougeron, C., Goldman, J.-P., Dart, A., Guelat, L., et Jeager, C. (2001). Influence de facteurs stylistiques, syntaxiques et lexicaux sur la réalisation de la liaison en français. *Paper presented at 8<sup>e</sup> TALN*, Tours.
- Francis, W. et Kucera, H. (1982). *Frequency Analysis of English Usage, Lexicon and Grammar*. Houghton Mifflin Company, Boston.
- Fries, C. (1945) *Teaching and Learning English as a Second Language*. Ann Arbor, University of Michigan Press.
- Gadet, F. (1989). *Le français ordinaire*. Paris: Colin.
- Gaillard, B., et Colignon, J.-P. (2005). *Toute l'orthographe*. Les dicos d'or de Bernard Pivot. Albin Michel.
- Gardner, H. (1993). *Frames of Mind: The Theory of Multiple Intelligences*. Britain. Fontana Press.
- Gardner, R.C. (1982). Language attitudes and language learning. In E. Bouchard Ryan & H. Giles, *Attitudes Towards Language Variation*, 132-147. Edward Arnold.
- Gardner, R.C. (2001). Language Learning Motivation: the Student, the Teacher, and the Researcher. *Texas Papers in Foreign Language Education*, 6, 1-18.
- Gass, S. et Lakshmanan, U. (1991). Accounting for interlanguage subject pronouns. *Second Language Research* 7, 181-203.
- Gass, S., et Mackey, A. (2002). Frequency effects and second language acquisition: A complex picture? *Studies in Second Language Acquisition*, 24(2), 249-260.
- Germain, C. (1993). *Évolution de l'enseignement des langues : 5000 ans d'histoire*. CLE international, Paris.
- Golato, P. S. (2002). Word parsing by late learning French-English bilinguals. *Applied Psycholinguistics*, 23 (3), 417-446.
- Goldsmith, J. A. (1976). *Autosegmental Phonology*. Ph.D. dissertation : MIT. New York : Garland Press, 1979.
- Gougenheim, G. (1938). *Le système grammatical de la langue française*. Paris : Bibliothèque du français moderne.
- Gouin, F. (1880). L'art d'enseigner et d'étudier les langues, Paris, G. Fichbacher, 589 p. [trad. angl. *The Art of teaching an Studying Languages*, London, 1892].
- Grammont, M. (1914). *La prononciation française*. Delagrave.

- Green, J. N., et Hintze, M.-A. (1990). *Variation and change in French linking phenomena*. Variation and change in French : Essays presented to Rebecca Posner on the occasion of her sixtieth birthday, (Eds.) J. N. Green and W. Ayres-Bennett, 61-88. London/New York : Routledge.
- Grevisse, M. (Eds.). (1980). *Le Bon Usage*. Gembloux : Duculot.
- Griessler, M. (2001) The effects of third language learning on second language proficiency: an Austrian example. *International Journal of Bilingual Education and Bilingualism*, 4 (1), 50-60.
- Hayek, F.A. (1952). *The Sensory Order: An Inquiry into the Foundations of Theoretical Psychology*. University Of Chicago Press 1999
- Hammarberg, B. (2009). *Processes in Third Language Acquisition*. Edinburgh: Edinburgh University Press.
- Hebb, D. O. (1949). *The Organization of the Behavior*, New York: Wiley
- Hieke, A. E. (1984). Linking as a marker of fluent speech. *Language and Speech*, 27 (4), 343-353.
- Hişmanoğlu, M. (2006). Current perspectives on pronunciation learning and teaching. *Journal of Language and Linguistic Studies*. 2 (1), 101-110  
<http://www.jlls.org/Issues/Volume%202/No.1/mhismanoglu.pdf>
- Hoff, E. (2002). Causes and consequences of SES-related differences in parent-to-child speech. In: M.H. Bornstein and R.H. Bradley (dir.), *Socioeconomic Status, Parenting, and Child Development* (pp. 147–160). Mahwah, NJ: Lawrence Erlbaum Associates.
- Howard, M. (2004). Sociolinguistic variation and second languages acquisition: A preliminary study of advanced learners of French. *SKY Journal of linguistics*, 17, 143-165.
- Howard, M. (2005). L'acquisition de la liaison en français langue seconde: Une analyse quantitative d'apprenants avancés en milieu guidé et en milieu naturel. *Corela*, Numéros spéciaux: Colloque AFLS.
- Howard, M. (2006). Variation in advanced French interlanguage: A comparison of three (socio) linguistic variables. *The Canadian Modern Language Review/ La Revue Canadienne des Langues Vivantes*, 62 (3), 379-400.
- Huot, F. (1978). *Difficultés phonétiques de l'acquisition du français langue seconde*. Québec : Centre international de recherches sur le bilinguisme.
- Huchon, M. (2002). *Histoire de la langue française*. Librairie Générale Française.
- Inglis, D. (2004) SIL Electronic Book Reviews 2004-004 - Usage-based models of language Usage-Based Models of Language, ed. by Michael Barlow and Suzanne Kemmer.  
 URL: <http://www.sil.org/silebr/2004/silebr2004-004>

- Itard, J.M. (1962). *The Wild Boy of Aveyron*. (L'enfant sauvage). New York: Wiley.
- Jelinek, F. (1976). Continuous speech recognition by statistical methods. *Proceedings of the IEEE*, 64 (4), 532–556
- Jones, R. (1997). Beyond “Listen and Repeat”: Pronunciation Teaching Materials and Theories of Second Language Acquisition. *System* 25 (1), 103-112.
- Juliano, C. et M. K. Tanenhaus (1993). Contingent frequency effects in syntactic ambiguity resolution. In *Proceedings of the 15<sup>th</sup> Annual Conference of the Cognitive Science Society, COGSCI-94*, Boulder (pp.593-598)
- Juliano, C. et Tanenhaus, M.K. (1994). A constraint-based lexicalist account of the subject/object attachment preference. *Journal of Psycholinguistic Research*, 23, 459-471.
- Kellerman, E. (1995). Crosslinguistic influence: Transfer to nowhere? *Annual review of Applied Linguistics*, 15, 125-150.
- Kjellin, O. (1999). *Accent Addition: Prosody and Perception Facilitates Second Language Learning*. In O. Fujimura, B. D. Joseph, & B. Palek (Eds.), *Proceedings of LP'98 (Linguistics and Phonetics Conference) at Ohio State University, Columbus, Ohio, September 1998 (Vol. 2, pp. 373-398)*. Prague: The Karolinum Press.  
<http://olle-kjellin.com/SpeechDoctor/ProcLP98.html>
- Klausenburger, J. (1974). Rules inversion, opacity, conspiracies: French liaison and elision. *Lingua*, 34 (2–4), 167–173.
- Klausenburger, J. (1976). French linking phenomena: a natural generative analysis. communication, *Summer Institute*. LSA : Oswego.
- Klausenburger, J. (1978). French linking phenomena: A natural generative analysis. *Language*, 54 (1), 21-40.
- Klausenburger, J. (1984). *French Liaison and Linguistic Theory*. Stuttgart: Steiner.
- Knoerr, H. (2000). Pratique intonative et utilisation d'un logiciel de visualisation dans un cours de prononciation en français langue seconde : une étude descriptive. *Revue canadienne de linguistique appliquée*, 3 (1-2), 123-138
- Knoerr, H., et Weinberg, A. (2006). L'enseignement de la prononciation en français langue seconde : de la cassette au cédérom. 2005 *La Revue canadienne des langues vivantes*, 61(3), 383-405.
- Krashen, S. (1979). Adult second language acquisition as post-critical period learning. *ITL: Review of Applied Linguistics*, 43, 39-52.
- Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. New York: Prentice-Hall.
- Krashen, S. D. (1985). *The Input Hypothesis*. London, Longman.

- Krashen, S., Long, M., et Scarcella, R. (1979). Age, rate, and eventual attainment in second language acquisition. *TESOL Quarterly*, 13, 573-582.
- Lado, R. (1957) *Linguistics across Cultures*. Ann Arbor, University of Michigan Press.
- Laeufer, C. (1985). *Language-specific and universal aspects of syllable structure and syllabification: Evidence from French and German*. Ph.D. Dissertation, Cornell University.
- Laeufer, C. (1994). French liaison and enchaînement. *The SECOL Review*, 18 (1), 1-30.
- Laks, B. (1983). Langage et pratiques sociales. Étude sociolinguistique d'un groupe d'adolescents. *Actes de la recherche en sciences sociales*, 46, 73-97.
- Laks, B. (2000). De la variation et des variantes : à propos du relâchement" *Linx*, 42, 21-29.
- Laks, B. (2005). La liaison et l'illusion. *Langages*, 158, 101-125.
- Lamb, S. (1999). *Pathways of the Brain: The Neurocognitive Basis of Language*. Benjamins.
- Langacker, R. W. (1987). *Foundations of Cognitive Grammar, vol.1: Theoretical Prerequisites*. Stanford, California: Stanford University Press.
- Langacker, R. W. (1991). *Foundations of Cognitive Grammar, vol.2: Descriptive application*. Stanford, California: Stanford University Press. Austin, Texas: University of Texas Press.
- Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. New York: Oxford University Press.
- Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, 25, 125-135.
- Larsen-Freemann D. et Long, M.H. (1991). *An Introduction to Second Language Acquisition Research*. London: Longman.
- Lauret, B. (2007). *Enseigner la prononciation du français : questions et outils*. Hachette.
- Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather and A. James (Eds.): *New Sounds 90: Proceedings of the Amsterdam Symposium on the Acquisition of Second Language Speech*. Amsterdam: University of Amsterdam, pp. 305-341.
- Lefrançois, P. (2005). How do university students solve linguistic problems? A Description of the Processes Leading to Errors. *L1-Educational Studies in Language and Literature*, 5 (3), 417-432.
- Léon, P. (1978). *Prononciation du français standard*. Montréal-Paris : Didier.
- Léon, P. (1992). *Phonétisme et Prononciations du français*. Paris: Éditions Fernand Nathan.
- Léon, P., et Léon, M. (1976). *Introduction à la phonétique corrective*. [Paris] : Hachette.

- Léon, P., et Tennant, J. (1990). "Bad French" and nice guys: A morphophonetic study. *The French Review*, 63 (5), 763-778.
- Lieven, E., Pine, J., et Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language*, 24, 187-220.
- Lightbown, P. (1983), Acquiring English L2 in Quebec classroom' in S. Felix and H. Wode (Eds.): *Language Development at the Crossroads*. Tübingen; Gunter Narr. pp. 151-175.
- Lightbown, P. (2003). SLA research in the classroom/SLA research for the classroom. *Language Learning Journal*, 28, 4-13
- Lightbown, P. M. et Spada, N. (2006). *How languages are learned* (3<sup>rd</sup> edition) Oxford: Oxford University Press.
- Lodge R. A., Armstrong, N., Ellis, Y. M. L., et Shelton J. F. (1997). *Exploring the French Language*. A Hodder Arnold Publication. pp. 95-100
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. *Handbook of Research On Second Language Acquisition*. W. Ritchie and T. Bhatia. New York, Academic.
- Long, M. H. et Robinson, P. (1998). "Focus on form: Theory, research and practice". In Doughty, Catherine; Williams, Jessica. *Focus on Form in Classroom Second Language Acquisition* (pp. 15-41). Cambridge: Cambridge University Press.
- Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics*, 15, 263-287.
- Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48, 183-218.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20, 51-81.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432
- Lyster, R. (2004). Research on form-focused instruction in immersion classrooms: Implications for theory and practice. *Journal of French Language Studies*, 14, 321-341. Cambridge University Press
- Lyster, R. et Ranta, L. (1997). Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, 19, 37-66.
- MacWhinney, B. (1978). The acquisition of morphophonology. *Monographs of the Society for Research in Child Development*, 43, Whole no. 1.
- MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), *Language acquisition: vol 1. Syntax and semantics*. pp. 73-136. Hillsdale, NJ: Lawrence Erlbaum.

- MacWhinney, B. (1987). Applying the competition model to bilingualism. *Applied Psycholinguistics*, 8, 315-327.
- MacWhinney, B. (1988). Competition and Teachability. In R. Schiefelbusch & M. Rice (Eds.), *The Teachability of Language*. pp. 63-104. New York: Cambridge University Press.
- MacWhinney, B. (1992). Transfer and competition in second language learning. In R.J. Harris (Ed.), *Cognitive Processing in Bilinguals* (pp. 371-390). Amsterdam: North Holland
- MacWhinney, B. (2004). A Unified Model of Language Acquisition. In J. Kroll & A. De Groot (Eds.) *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford University Press.
- Malécot, A. (1954). The pronunciation of numerals in French. *The French Review*, 27 (4), 287- 297.
- Malécot, A. (1975). French Liaison as a Function of Grammatical, Phonetic and Paralinguistic Variables. *Phonetica*, 32, 161-179.
- Malécot, A. (1977). The effect of linguistic and paralinguistic variables on the elision of the French mute e. *Phonetica*, 33, 93-112.
- Mallet, G.-M. (2008). *La liaison en français : descriptions et analyses dans le corpus PFC*. Thèse de doctorat. Université Paris Ouest Nanterre La Défense. École Doctorale 139 : Connaissance, Langage, Modalisation. Laboratoire. pp. 328
- Markham, D. J., et Nagano-Madsen, Y. (1996). Input modality effects in foreign accent. In: *Proc. of the 1996 International Conference on Spoken Language Processing 1996*, pp. 1473–1476.
- Martinon, P. (1913). *Comment on prononce le français. Traité complet de prononciation pratique avec les noms propres et les noms étrangers*. Paris : Larousse.
- Mastromonaco, S. M. (2000). *Liaison in French as a second language*. Dissertation Abstracts International. (UMI No. DANQ41234).
- Matthews, D., Lieven E. V., Theakston A. L., et Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, 20, 121-136.
- Moisset, C. (2000). Variable liaison in Parisian French. *Dissertation Abstracts International, A: The Humanities and Social Sciences*, 61 (03), 965-A-966-A.
- Monnin, J., Loevenbruck, et H., Beckman, M. E. (2007). The influence of frequency on word-initial obstruent acquisition in *Hexagonal French 16th International Congress of Phonetic Sciences, ICPHS 2007*, Sarrebruck : Germany (2007).
- Morin, Y. C. (2003) [1998]. Remarks on prenominal liaison consonants in French. *Living on the Edge, 28 Papers in Honour of Jonathan Kaye*, Stefan Ploch (Ed.), Mouton de Gruyter.

- Morin, Y.-C. (2005). La liaison relève-t-elle d'une tendance à éviter les hiatus? Réflexions sur son évolution historique. *Langages*, 158, 8-23.
- Morin, Y.-C., et Kaye, J. (1982). The syntactic bases for French liaison, *Journal of Linguistics*, 18, 291-330.
- Morin, Y.-C., Langlois, M-C. et Varin, M-È. (1990). Tensing of word-final [ɔ] to [o] in French : the phonologization of a morphophonological rule. *Romance Philology*, 43, 507-528.
- Mougeon, R., T. Nadasdi, et K. Rehner, (2010) *The Sociolinguistic Competence of Immersion Students*. Clevedon, UK: Multilingual Matters.
- Nagamine, T. (2002). An experimental study on the teachability and learnability of English intonational aspect: Acoustic Analysis on F0 and native-speaker judgment task. *Journal of Language and Linguistics*, 1(4), 1-16.
- Nardy, A. (2003). *Production et jugement d'acceptabilité entre 2 et 6 ans : aspects psycholinguistiques et sociolinguistiques de l'acquisition des liaisons*, Diplôme d'Études Approfondies, Université Grenoble 3.
- Neufeld, G. (1977). Language learning ability in adults: a study on the acquisition of prosodic and articulatory features. *Working Papers on Bilingualism*, 12, 45-60.
- Neufeld, G. (1978). On the acquisition of prosodic and articulatory features in adult language learning. *Canadian Modern Language Review* 34 (2), 163-74.
- Neufeld, G. (1979). Vers une théorie de la capacité d'apprentissage linguistique. *Encrages, n° spécial de linguistique appliquée*. Université de Paris VIII.
- Neufeld, G. (1980). On the adult's ability to acquire phonology. *TESOL Quarterly*, 14, 285-298
- New, B., C. Pallier, L. Ferrand, et R. Matos (2001). Une base de données lexicales du français contemporain sur Internet : LEXIQUE. *L'Année Psychologique*, 101, 447-462.
- New, B., Brysbaert, M., Veronis, J., et Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics*, 28(4), 661-677.
- Nyrop, K. (1934). *Manuel phonétique du français parlé*. Copenhague : Nordiske.
- Nyrop, K. (1935). *Grammaire historique de la langue française*, vol. I: *Histoire générale de la langue française, Phonétique historique*, 4th edition Copenhague: Gyldendalske Boghandel, Nordisk Forlag.
- Odlin, T. (1989). *Language Transfer*. Cambridge : Cambridge University Press.
- Ollivier, J., et Beaudoin, M. (2010). *Grammaire française, 5<sup>e</sup> édition*, Toronto (Ontario), Nelson Education.



- Ostiguy, L., et Sarrasin, R. (1985). *Phonétique comparée du français et de l'anglais nord-américains*. Trois-Rivières, Québec : Les Éditions du Réseau U Inc.
- Ovtcharov, V., Cobb, T. et Halter, R. (2006). La richesse lexicale des productions orales : mesure fiable du niveau de compétence langagière. *Canadian Modern Language Review*, 63 (1), 107-125.
- Özçelik, N. (2008). Problèmes de prononciation des étudiants turcs en français. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi* (H. U. Journal of Education) 34, 204-217.  
<http://193.140.216.63/200834NURTEN%20%C3%96Z%C3%87ELIK.pdf>
- Pagliano C., et Laks, B. Problématiques de la liaison dans l'analyse d'un corpus de français oral actuel. Université Paris X. Retrieved : June 25, 2008, from  
[http://colloqueff.ens-lsh.fr/pdf/Laks\\_Bernard\\_Pagliano\\_Claudine.pdf](http://colloqueff.ens-lsh.fr/pdf/Laks_Bernard_Pagliano_Claudine.pdf)
- Pallaud, B., et Savelli, M. (2001). L'oral enfantin. Quelques précautions pour l'évaluer. *Revue française de linguistique appliquée*, 6 (1), 121-136.
- Parrish, B. et E. Tarone (1988). Article usage in interlanguage: A study in task-related variability, *Language Learning*, 38, 21-44.
- Passy, P. (1892). *Les sons du français, 3e édition*. Paris, Firmin-Didot.
- Pernot, N. (1937). Quelques notes sur la liaison en français: Liaison et enchaînement. *The Modern Language Journal*, 21 (5), 333-338.
- Le Petit Robert (2002) - *Dictionnaire de la langue française*. 2592 pages, 2002 éd.
- Peyrollaz, M., et de Tovar, B. M. L. (1954). *Manuel de phonétique et de diction françaises* (pp.173-198). Larousse : Paris - 6e.
- Pierrehumbert, J. (2001). Stochastic Phonology. *GLOT*, 5 (6), 195-207.
- Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven and N. Warner (Eds.), *Laboratory Phonology VII*, Mouton de Gruyter, 101-140.
- Pienemann, M., et Johnston, M. (1986). An acquisition based procedure for second language assessment (ESL). *Australian Review of Applied Linguistics*, 9 (1), 92-122.
- Preston, D. (2000). Three kinds of sociolinguistics and SLA: A psycholinguistic perspective. In F. M. B. Swierzbin, M. Anderson, C. Klee, & E. Tarone (Eds.), *Social and cognitive factors in second language acquisition: Selected proceedings of the 1999 Second Language Research Forum* (pp. 3-30). Somerville, MA: Cascadilla.
- Purcell, E., et Suter, R. (1980). Predictors of pronunciation accuracy: A reexamination. *Language Learning*, 30 (2), 271-87.
- Puren, C. (1988): *Histoire des méthodologies de l'enseignement des langues*, Paris, Nathan-Clé International, Col. DLE.

- Rand, D., et Sankoff, D. (1990). *GoldVarb: A Variable Rule Application for the Macintosh*. Montréal : Centre de recherches mathématiques, Université de Montréal.
- Ranson, D. (2008). La liaison variable dans un corpus du français méridional : L'importance relative de la fonction grammaticale. Congrès Mondial de Linguistique Française - CMLF'08. Durand J. Habert B., Laks B. (Eds.). ISBN 978-2-7598-0358-3, Paris, 2008, Institut de Linguistique Française. Phonétique, phonologie et interfaces. DOI 10.1051/cmlf08279  
[www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08279.pdf](http://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08279.pdf)
- Rehner, K. (2010a). The use/non-use of *ne* in the spoken French of university-level learners of French as a second language in the Canadian context. *French Language Studies*, 20, 289-311.
- Rehner, K. (2010b). The sociolinguistic competence of former immersion students at the post-secondary level: the case of lexical variation. *International Journal of Bilingual Education and Bilingualism*, 1-17, iFirst article.
- Rehner, K. et Mougeon, R. (1999). Variation in the spoken French of immersion students: To *ne* or not to *ne*, that is the sociolinguistic question. *Canadian Modern Language Review*, 56, 124-54.
- Rivenc, P., (1962 et 1964). Voix et images de France (1<sup>er</sup> et 2<sup>ème</sup> degrés). Cours et méthode audiovisuelle d'enseignement du français aux étrangers, réalisée par le CREDIF, sous la direction de P. Rivenc et P. Guberina, Paris, Didier, 4 manuels, 32 films fixes, 52 bandes magnétiques enregistrées.
- Rodríguez Seara, A. (2001). L'évolution des méthodologies dans l'enseignement du français langue étrangère depuis la méthodologie traditionnelle jusqu'à nos jours. Centro Asociado de la UNED de Tudela - Espagne. [http://www.uned.es/ca-tudela/revista/n001/art\\_8.htm](http://www.uned.es/ca-tudela/revista/n001/art_8.htm)
- Rumelhart, D. E., et McClelland, J. L. (1986). On learning the past tense of English verbs. In J. L. McClelland et D. E. Rumelhart (Eds.), *Parallel distributed processing: explorations in the microstructure of cognition*. Cambridge: MIT Press.
- Rymer, R. (1993). *Genie A Scientific Tragedy*. Harper Perennial.
- Sachs, J., B. Bard et Johnson, M. (1981). Language learning with restricted input: Case studies of two hearing children of deaf parent. *Applied Psycholinguistics*, 2 (1), 33-54.
- Saffran, J.R., Newport, E.L., et Aslin, R.N. (1996) Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928.
- Salcedo, C. (2002). *The effects of songs in the foreign language classroom on Text recall and involuntary mental rehearsal*. Ph.D Dissertation. 1-169  
[http://etd.lsu.edu/docs/available/etd-1111102-204823/unrestricted/Salcedo\\_dis.pdf](http://etd.lsu.edu/docs/available/etd-1111102-204823/unrestricted/Salcedo_dis.pdf)
- Sampson, R. (2001). Liaison, nasal vowels and productivity. *Journal of French Language Studies*, 11 (2), 241-258.

- Sauders, G. (2007). Vowel deletion and resyllabification in French speech: a source of perceptual difficulty for American learners of French. *Phonetics Teaching & Learning Conference*. UCL August 24-26, 2007  
[http://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper\\_26e.pdf](http://www.phon.ucl.ac.uk/ptlc/proceedings/ptlcpaper_26e.pdf)
- Schane, S. A. (1965). *The phonological and morphological structure of French*, Ph.D. dissertation : MIT Press.
- Schmidt, R. (2001). *Attention, cognition and second language instruction*. P. Robinson. Cambridge, Cambridge University Press: 3-32.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 17-46.
- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage Pragmatics* (pp. 21-42). Oxford University Press.
- Selkirk, E. O. (1974). French liaison and the X notation. *Linguistic Inquiry*, 5, 573-590.
- Selkirk, E. (1972). *The phrase phonology of English and French*, Ph. D. Dissertation, MIT, New York: Garland.
- Sergeeva, E., et Chevrot, J.-P. (2008). Influence de la fréquence sur la production et l'acquisition des formes verbales du français: le cas des russophones adultes. *Congrès Mondial de Linguistique Française, Paris, France, 2008*. Durand J. Habert B., Laks B. (éds.). Publié en ligne le 9 juillet 2008.
- Skinner, B. F. (1957). *Verbal Behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Smith, A. (1998). French variable liaison: A proposed simplification. *Francophonie*, 17, 11-14.
- Southworth, M.-J. (1970). French words in H-. *The French Review*, 44 (1), 63-71.
- Spada, N. (2007). Chapter 18. Communicative language teaching: Current status and future prospects. In J. Cummins & C. Davison (Eds.), *Kluwer handbook of English language teaching*. Amsterdam: Kluwer Publications.
- Spada, N. et P.M. Lightbown (1993). Instruction and the development of questions in L2 classrooms. *Studies in Second Language Acquisition*, 15 (2), 205-224.
- Spinelli, E., et Meunier, F. (2005). The cognitive treatment of liaison in the recognition of connected speech. *Langages*, 158, 79-88.
- Spitzer, L. (1941). French pataquès; cuirs. *Language*, 17 (3), 253-255.
- Sprenger-Charolles, L., Siegel, L. S., et Bonnet, P. (1998). Reading and spelling acquisition in French: The role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology*, 68 (2), 134-165.

- Sprenger-Charolles, L., et al. (2003). Development of Phonological and Orthographic Processing in Reading Aloud, in Silent Reading and in Spelling: A Four Year Longitudinal Study. *Journal of Experimental Child Psychology*, 84, 194-217.
- Stapp, Y. (1999). Neural plasticity and the issue of mimicry tasks in L2 pronunciation studies. *TESL-EJ*, 3 (4), 1-19.
- Steele, J. (2001). Ultimate attainment in L2 syllabification. *McGill Working Papers in Linguistics / Cahiers linguistiques de McGill*, 15 (2), 113-26.
- Stridfeldt, M. (2002). La perception du français par les apprenants suédophones. *Romansk Forum. XV Skandinaviske romanistkongress*, 16, 12-17.
- Syrovatskaja, R. (2008). Le rôle de la phonétique articulatoire et corrective dans l'enseignement du français. p.1-3  
<http://ressources-cla.univfcomte.fr/gerflint/Baltique2/phonetiquearticulatoire.pdf>
- Suter, R. W. (1976). Predictors of pronunciation accuracy in second language learning. *Language Learning*, 26, 233-253.
- Swain, M. (1985). Communicative competence: some roles of comprehensible input and comprehensible output in its development. *Input in Second Language Acquisition*. S. M. Gass and C. Madden. Rowley, Mass., Newbury House.
- Tajika, H. (1999). *Variable patterns of tense/aspect marking in interlanguage*. Unpublished dissertation, University of Minnesota, Minneapolis.
- Taraban, R., et Roark, B. (1996). Competition in learning language-based categories. *Applied Psycholinguistics*, 17, 125-148
- Tardif, C., et d'Anglejan, A. (1981). Les erreurs du français langue seconde et leurs effets sur la communication orale. *The Canadian Modern Language Review / La Revue Canadienne des Langues Vivantes*, 37, 706-723.
- Tarone, E. (1979). Interlanguage as chameleon, *Language Learning* 29, 181-191.
- Tarone, E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax, *Language Learning*, 35, 373-403.
- Tarone, E. (1988). *Variation in Interlanguage*. London: Edward Arnold.
- Tarone, E. (1998). A sociolinguistic perspective on an SLA theory of mind. *Studia Anglica Posnaniensia*, 33, 431-444.
- Tarone, E. (2002). Frequency effects, noticing, and creativity: Factors in a variationist interlanguage framework, *Studies in Second Language Acquisition*, 24, 287-296.
- Thomas, A. (1998). La liaison et son enseignement : des modèles orthoépiques à la réalité linguistique. *La Revue canadienne des langues vivantes*, 54 (4), 543-552.

- Thomas, A. (2002). La variation phonétique en français langue seconde au niveau universitaire avancé. *Aile: Acquisition et Interaction en Langue Étrangère*, 17, 101-121
- Thomas, A. (2004). Phonetic norm versus usage in advanced French as a second language. *IRAL*, 42 (4), 365-382.
- Thomas, A. (2008). La mesure des progrès lexicaux en FL2 avancé. *Congrès Mondial de Linguistique Française - CMLF'08*. Durand J. Habert B., Laks B. (Eds.) 587-597, Paris, 2008, Institut de Linguistique Française.
- Tomasello, M. (1992) *First Verbs: A Case Study in Early Grammatical Development*. Cambridge University Press.
- Tomasello, M. (2000a). The item-based nature of children's early syntactic development. *Trends in Cognitive Sciences*, 4, 156-163.
- Tomasello, M. (2000b). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11 (1/2), 61-82.
- Tomasello, M. (2003). *Constructing a Language: a Usage-Based Theory of Language Acquisition*. Cambridge [Massachusetts]: Harvard University Press.
- Tomasello, M. et Brooks, P. (1998). Young children's earliest transitive and intransitive constructions. *Cognitive Linguistics*, 9, 379-395.
- Tomasello, M., Akhtar, N., Dodson, K., et Rekau, L. (1997). Differential productivity in young children's use of nouns and verbs. *Journal of Child Language*, 24, 373-387
- Tomé, M. (2003). Les ressources Internet pour la phonétique du Français Langue Etrangère. *VIème Congrès International de Linguistique Française*, Universidad de Granada. [article en ligne : <http://www3.uni.es/dp/dfm/fenet/activi4.htm#ressourcesPhonetique>]
- Trahey M., et White L. (1993). Positive evidence and preemption in the second language classroom. *Studies in Second Language Acquisition*, 15, 181-204.
- Tranel, B. (1981). *Concreteness in Generative Phonology*. Evidence from French, Berkeley: University of California Press.
- Trubetzkoy, N. S. (1939). *Principes de phonologie*. Paris, Klincksieck.
- Tummers, J., K. Heylen et D. Geeraerts (2005). Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory*, 1, 225-261
- Valdman, A. (1993). *Bien Entendu ! Introduction à la prononciation française*. Englewood, Cliffs. NJ : Prentice Hall.
- Van Ameringen, A., et Cedergren, H. J. (1981). Observations sur la liaison en français de Montréal. In D. Sankoff et H. Cedergren (reds.) *Variation omnibus*, (pp. 141-149). Edmonton: Linguistics Research Inc.

- VanPatten, B. (Ed.). 2004. *Processing Instruction: Theory, Research and Commentary*. Lawrence Erlbaum Associates, NJ.
- Vaugelas, C. de Fabre. (1647). *Remarques sur la langue françoise*. Paris : Droz.
- Vendelin, I., et Peperkamp, S. (2005). *The influence of orthography on loanwords adaptations*. Paper presented at 7èmes Rencontres Internationales du Réseau Français de Phonologie, Université d'Aix-en-Provence.
- Venkatagiri, H. S., et Levis, J. (2007). Metaphonological knowledge and comprehensibility: An exploratory study. *Language Awareness*, 16, 263-277.
- Verdelhan, M. (2002) *Le français de scolarisation : pour une didactique réaliste*, Paris, PUF.
- Vygotsky, L.S. (1978). *Mind in Society*. Cambridge Mass.: Harvard University Press.
- Wagner-Gough, J. et Hatch, E. (1975). The importance of input data in second language acquisition studies. *Language Learning*, 25, 297-308.
- Walker, D. C. (1984). *The Pronunciation of Canadian French*. Ottawa: University of Ottawa Press.
- Watson, J. B. (1924). *Behaviorism*. New York: Norton.
- Wauquier-Gravelines, S. (2002). Statistical learning or phonological generalisation: The case of acquisition of liaison in French. *Special Session « Acquisition » of Tenth Phonology Meeting of Manchester*, 23 mai 2002.
- Wauquier-Gravelines, S. (2003). Troncation et reduplication, peut-on parler de gabarits dans le lexique précoce ? In Fradin, B., Dal, G., Hathout, N., Kerleroux, F., Roché, M. & Plénat, M. (éds), *Les unités morphologiques, Silexicales*, 3, 220-229, Université de Lille III, 2003.
- Wauquier-Gravelines, S., et Braud, V. (2005). Proto-déterminant et acquisition de la liaison obligatoire en français. *Langages*, 158, 53-65.
- Yates, K. (2003). *Teaching Linguistic Mimicry to Improve Second Language Pronunciation*. Thesis Prepared for the Degree of Master of Arts. University of North Texas.
- Young, R. (1991). *Variation in Interlanguage Morphology*. New York. Peter Lang.
- Wang, X. (2007). Effects of Length Of Residence And Speech Activities On Degree Of Foreign Accent. *ICPhS - ID 1114 -Saarbrücken*, 6-10 August 2007  
<http://www.icphs2007.de/conference/Papers/1114/1114.pdf>
- Zobl, H. (1982). A direction for contrastive analysis: the comparative study of developmental sequences. *TESOL Quarterly*, 16 (2), 169-183.

## Nadine de Moras

**Ph.D. French Studies** with concentration in Applied Linguistics. Title of thesis : Acquisition de la liaison et de l'enchaînement en français L2, le rôle de la fréquence. UWO, 2011

**M. A. French Studies** with concentration in Applied Linguistics: Second Language (French) Acquisition and Teaching UWO, 1997

**B.A. English and Russian:** English for Business and Economics with a minor in Russian Paul Valéry University, Montpellier, France, 1994

## PUBLICATIONS

N. de Moras, (2010). *Guide d'écriture*. 1<sup>st</sup> Edition. Thomson et Nelson. 256 p.

V. Prat, F. Poiré et N. de Moras, (2004). De la salle de classe au laboratoire et vice versa: L'amélioration de la compréhension globale et fine grâce à la réutilisation du matériel audio-visuel sur CAN-8. *Distances*, 6 (2), 69-92.

G. Holmes et N. de Moras (1997). A French Language Grammar Analyser: What Use for Anglophone Students? In *FLEAT III: Foreign Language Education and Technology*. Ed. P. Liddell. University of Victoria. (pp. 91-106)

## EMPLOYMENT HISTORY

<b>FULL-TIME LECTURER</b> Brescia University College	2007-2011
<b>PART-TIME LECTURER</b> Affiliated UWO University Colleges	2005-2007
<b>FULL-TIME LECTURER</b> Simon Fraser University (SFU)	2002-2004
<b>PART-TIME LECTURER</b> main campus and Affiliated University Colleges	1995-2002

## RECOGNITION OF TEACHING ACHIEVEMENT

- **Dean's honour roll of teaching** Brescia University College, 2009-2011
- **Teaching Award** nominee Brescia University College, 2008
- **Dean's honour roll of teaching** November 8, 2002
- **Graduate Student Teaching Award** nominee UWO, 2000-2002
- **Dean's honour roll of teaching** 1999-2000