11-29-2023 10:30 AM

Differentiated Performance in Native and Instructed Nonnative Spanish: The impact of task demands and individual differences during performance with grammatical gender

Martha Black,

Supervisor: Bruhn de Garavito, Joyce, *The University of Western Ontario*
Co-Supervisor: Herment, Sophie, *Aix Marseille Université*
A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Hispanic Studies

Abstract

Adult language learners demonstrate extensive variation and are believed to rely largely on explicit knowledge and declarative memory, directly impacting how target language input is processed, represented, and retrieved. The overarching aim of this thesis is to identify and better understand the factors that are most important for adult language acquisition by examining how linguistic features, task demands, and individual learner differences may impact performance with Spanish grammatical gender. This research seeks to draw principled conclusions about what knowledge types and memory systems language users exploit as proficiency develops. 115 language users of Spanish from diverse language backgrounds, including native speakers (n=25) and late/adult instructed Spanish learners (n=90), completed a language learner profile questionnaire, a Spanish proficiency test, four experimental tasks with strategically manipulated conditions, and a metalinguistic awareness exit survey. Findings indicate enhanced late learner accuracy with frequent, grammatical, and masculine noun tokens and slightly enhanced native speaker accuracy with high-frequency tokens. Intermediate and beginner learner performance was found to be enhanced on self-paced and written tasks whereas advanced learners and native speakers performed better on speeded tasks and showed no stimuli modality effects. Individual differences in Spanish proficiency, metalinguistic awareness, and the Ideal L2 Self component of motivation produced the largest effect sizes among late learners. A slight typological multilingual learner advantage was also found that produced the greatest learner advantage on tasks conditioning online language processing. This thesis contributes to the domain of adult language acquisition by providing evidence that as global proficiency in the target language develops, qualitative patterns of sensitivity to linguistic and task features become more native-like as do quantitative measures of performance. Findings suggest that at lower levels of proficiency, learners strategically exploit their explicit linguistic knowledge to compensate for deficits in their developing implicit linguistic system. This research further contributes to our understanding of the individual differences that impact performance and makes a novel contribution to the field of multilingualism by elucidating the nature of the multilingual advantage for language learning. Findings show that advanced proficiency late learners are able to mirror both quantitative and qualitative native speaker norms of performance. Pedagogical implications are also discussed.

Keywords

adult language acquisition, instructed language acquisition, Spanish, morphosyntax, grammatical gender, multilingualism, individual differences, task effects, implicit/explicit language processing.

Summary for Lay Audience

This thesis looks at how linguistic qualities, task requirements, and learner differences affect performance with Spanish grammatical gender in order to better understand the factors which influence adult language acquisition. A questionnaire, a Spanish proficiency test, experimental tasks, and an exit survey on metalinguistic awareness were all completed by 115 Spanish language users from various backgrounds. The findings indicated that high-frequency tokens somewhat improved the accuracy of native speakers, while frequent, grammatical, and masculine noun tokens improved the accuracy of late learners. On written and self-paced tasks, learners who were beginner or intermediate performed better while native speakers and advanced learners did better on tasks that were speeded up. The greatest effect sizes among late learners were found in individual differences in Spanish proficiency, metalinguistic awareness, and the Ideal L2 Self component of motivation. This study adds to our knowledge of individual differences and multilingualism and suggests that late learners of advanced proficiency can closely resemble the performance standards of native speakers.

Resumen *(Spanish)*

Los aprendices adultos de idiomas demuestran una amplia variación y dependen en gran medida del conocimiento explícito y la memoria declarativa, lo que afecta directamente la forma en que se procesa, representa y recupera la entrada del lenguaje. El objetivo general de esta tesis es identificar y comprender mejor los factores que son más importantes para la adquisición del lenguaje en los adultos mediante el análisis de cómo las características lingüísticas, las demandas de tareas, y las diferencias individuales de los alumnos pueden afectar el rendimiento con el género gramatical español. Buscamos sacar conclusiones basadas en principios sobre qué tipos de conocimiento y sistemas de memoria explotan los usuarios del idioma a medida que se desarrolla la competencia. 115 usuarios de español de diversos orígenes lingüísticos, incluidos hablantes nativos (n = 25) y aprendices adultos de español (n = 90) del contexto instruido, completaron un cuestionario de perfil de estudiante de idiomas (“language learner profile questionnaire”), una prueba de competencia en español, cuatro tareas experimentales con condiciones manipuladas estratégicamente, y una encuesta de salida de conciencia metalingüística. Los resultados indican una preferencia de los aprendices tardíos por los tokens de sustantivos frecuentes, gramaticales y masculinos y una ligera preferencia de los hablantes nativos por los tokens de alta frecuencia. Se encontró que el rendimiento de los aprendices intermedios y principiantes fue mejor en las tareas escritas y las tareas a su propio ritmo (“self-paced”), mientras que los aprendices avanzados y los hablantes nativos se desempeñaron mejor en las tareas aceleradas (“speeded”) y no mostraron efectos de modalidad de estímulo. El dominio del español produjo el efecto más diferenciador entre los aprendices tardíos, seguido por la conciencia metalingüística y el componente de motivación del ‘yo ideal en la L2’ (“Ideal L2 Self”). Se encontró una ligera ventaja tipológica del multilingüismo que produjo la mayor ventaja para los aprendices en las tareas que condicionan el procesamiento del lenguaje en tiempo real (“online language processing”). Esta tesis contribuye al dominio de la adquisición del lenguaje en adultos al proporcionar evidencia de que a medida que se desarrolla la competencia global en el idioma de destino, los patrones cualitativos de sensibilidad a las funciones lingüísticas y de tareas se vuelven más parecidos a los nativos, al igual que las medidas cuantitativas de desempeño. Los resultados sugieren que en los niveles más bajos de competencia, los aprendices explotan estratégicamente su conocimiento lingüístico explícito para compensar las deficiencias en su sistema lingüístico implícito en desarrollo. Esta investigación contribuye a nuestra comprensión de las diferencias individuales que afectan el rendimiento y hace una contribución novedosa al campo del multilingüismo al dilucidar la naturaleza de la ventaja multilingüe. Los hallazgos muestran que los aprendices tardíos (adultos) de competencia lingüística avanzada en español pueden reflejar las normas de desempeño tanto cuantitativas como cualitativas de los hablantes nativos. También se discuten las implicaciones pedagógicas.

Palabras clave

la adquisición adulta del lenguaje, la adquisición instruida del lenguaje, el español, la morfosintaxis, el género gramatical, el multilingüismo, las diferencias individuales, los efectos de tarea, el procesamiento lingüístico implícito/explícito.

Résumé *(French)*

Les apprenants adultes de la langue présentent des variations importantes et sont censés s'appuyer en grande partie sur les connaissances explicites et la mémoire déclarative, ce qui a un impact direct sur la façon dont l'entrée de la langue cible est traitée, représentée et récupérée. L'objectif principal de cette thèse est d'identifier et de mieux comprendre les facteurs les plus importants pour l'acquisition du langage chez les adultes en examinant comment les caractéristiques linguistiques, les exigences des tâches et les différences individuelles des apprenants peuvent avoir un impact sur les performances liées au genre grammatical espagnol. Nous cherchons à tirer des conclusions de principe sur les types de connaissances et les systèmes de mémoire que les utilisateurs du langage exploitent au fur et à mesure que la compétence se développe. 115 locuteurs d'espagnol issus de divers horizons linguistiques, des locuteurs natifs (n = 25) et des apprenants d'espagnol tardifs / adultes (n = 90), ont rempli un questionnaire sur le profil de l'apprenant en langue, un test de compétences en espagnol, quatre tâches expérimentales avec des conditions stratégiquement manipulées et une enquête de sensibilisation métalinguistique à la fin. Les résultats indiquent une préférence des apprenants tardifs pour les occurrences de noms fréquents, grammaticaux et masculins et une légère préférence du locuteur natif pour les occurrences à haute fréquence. Les performances des apprenants intermédiaires et débutants se sont avérées améliorées sur les tâches auto-rythmées et écrites, tandis que les apprenants avancés et les locuteurs natifs ont obtenu de meilleurs résultats sur les tâches accélérées et n'ont montré aucun effet de modalité de stimuli. La maîtrise de l'espagnol a produit l'effet le plus différenciateur entre les apprenants tardifs, suivie de la conscience métalinguistique et de la composante de motivation du soi idéal en L2 (« Ideal L2 Self »). Les apprenants multilingues ont montré de meilleurs résultats pour les tâches conditionnant le traitement du langage en temp réel. Cette thèse contribue au domaine de l'acquisition du langage chez l'adulte en apportant la preuve qu'à mesure que la maîtrise globale de la langue cible se développe, les schémas qualitatifs de sensibilité aux caractéristiques linguistiques et aux tâches deviennent plus natifs, tout comme les mesures quantitatives de la performance. Les résultats suggèrent qu'à des niveaux de compétence inférieurs, les apprenants exploitent stratégiquement leurs connaissances linguistiques explicites pour compenser les déficits de leur système linguistique implicite en développement. Cette recherche contribue davantage à notre compréhension des différences individuelles qui ont un impact sur la performance et apporte une nouvelle contribution au domaine du multilinguisme en élucidant la nature de l'avantage multilingue. Les résultats montrent que les apprenants tardifs à compétences avancées sont capables de refléter les normes de performance quantitatives et qualitatives des locuteurs natifs. Les implications pédagogiques sont également discutées.

Mots clés

acquisition tardive du langage, acquisition du langage en contexte formel, espagnol, morphosyntaxe, genre grammatical, multilinguisme, différences individuelles, effets de tâche, traitement implicite/explicite du langage.

Acknowledgements

I received much help from others throughout the process of my doctoral education—from professors, administrative staff, collaborators, and close friends and family. Like anything truly significant in life, we never go it alone, and my graduate education is no exception.

First and foremost, I would like to thank my kind, understanding, tolerant, inquisitive, and inspiringly authentic thesis supervisor, Dr. Joyce Bruhn de Garavito, without whose academic, emotional, and financial support the quality of this thesis would not have been the same. Her attention to detail, the quality and honesty of her feedback, and her kindheartedness have been unmatched in my graduate experiences at Western University. She was so willing to pivot our meetings and the design of my research when the COVID-19 pandemic hit and everything moved online, and she has been equally understanding of my life changes that have compelled me to move to France and start a teaching career here. Furthermore, Joyce has continually fueled and supported my passion for operationalizing the all-important connection between language acquisition research and foreign language pedagogy in the classroom. I plan to dedicate the rest of my academic career to continuing to bridge this divide. Finally, Joyce's conviction that adults *are*, in fact, capable of fully acquiring language is one that continually inspires me to keep working towards this objective of bridging the gap—it *is* possible, and we *can* do better to support adults in their language learning journey. Gracias, Joyce.

I also owe much gratitude to the Department of Languages and Cultures at Western University for their continued financial support of my education and research as well as for their generous understanding and flexibility allowing me to complete my studies remotely while starting out a new teaching career in France.

I am so thankful to Dr. Sophie Herment, my thesis co-supervisor, and the Laboratoire Parole et Langage (LPL) at Aix Marseille Université in France, who kindly opened their doors to me, accepting me into their research community. Sophie's kindness and openness to accept me into the stimulating community of linguists at the lab has helped me to establish myself academically at Aix Marseille Université here in France and ultimately propelled me to finish my doctoral research remotely with a dedicated lab space for me to work and meaningful connections with local colleagues in the field, all of which has provided me with the all-important sense of connection and belonging in my new home. Thank you, Sophie.

I would also like to thank Frank Boers, in the Faculty of Education at Western University, for his help as my initial supervisor on the early stages of my thesis and ultimately as second reader, as well as for his support as the supervisor of my original course design. His applied linguistics and education backgrounds were enormously helpful when designing my course that aimed to connect acquisition theory to pedagogical practice and has ultimately informed the direction of my doctoral research. Thank you, Frank.

Sylvia Kontra, Graduate Program Assistant in the Department of Languages and Cultures at Western University, has been immensely helpful throughout my graduate trajectory at Western, continually keeping me informed and organized and also providing me with encouragement and support. Her ability to balance a demanding and complex workload with compassion and consideration for the graduate students of the department, coupled with her attention to detail and her understanding of our needs, is truly exemplary and one that I hope to emulate in my future career path with students. Thank you, Sylvia.

I also must thank Dr. Yasaman Rafat for taking me under her wing early on on several projects and helping me to publish my first academic peer-reviewed article in 2020. Our collaborations continue today with her newly established Second Language Acquisition and Multilingual (SLAM) Lab, and I look forward to our collaborations to come. Thank you, Yasaman.

I owe much thanks to Dr. Olga Tararova for hiring me as a research assistant to work under her direction in developing and carrying out a multi-year research project on multilingualism and adult language acquisition. This continual project and funding have helped me to refine my research skills and gain extensive experience above and beyond my thesis research in the design, implementation, and data analysis and interpretation that fuel high-quality empirical research in language acquisition. As a young scholar herself, Olga understands how starting out as a new researcher and academic is challenging and requires a team of support, which she has so generously provided to me and many others. Thank you, Olga.

I would also like to thank Dr. Ana García-Allén, Department Chair in Languages and Cultures at Western University, for her encouragement and support of language acquisition research in the

Department by facilitating the connection between undergraduate language students and graduate student researchers. Gathering a rather large sample of language learners would not have been possible without her support and guidance in addition to her continued support of research that works to bridge the gap between language acquisition and language pedagogy. Furthermore, my experiences teaching beginner and intermediate Spanish courses under her direction have been a truly formative opportunity that has informed by research objectives in studying how adults learn languages in the instructed context with the ultimate objective of better supporting adult language learning in the classroom. Gracias, Ana.

I am eternally grateful to Dr. Patricia Fagan at the University of Alaska Anchorage (UAA), where I taught Spanish as an adjunct instructor under her direction between 2015 and 2018. Although teaching as an adjunct instructor firmly places one at the periphery of the academic community, Patricia went out of her way to bring me into the fold by including me in department meetings, activities, and decisions, and she encouraged me to continue my education by pursuing a doctoral degree. Those three years I spent teaching Spanish to adults in Alaska helped me to hone my research interests in adult language acquisition in the instructed context. During my tenure at UAA, I realized that, as language educators, we must make it our *imperative* to consider the implications of language acquisition research for language teaching in order to better support adult language learning outcomes in the classroom, and it is my hope that my doctoral thesis makes a novel and meaningful contribution toward this ultimate objective. Thank you, Patty.

I am also immensely grateful to Francis Alexander Frías Lora, my ex-husband and dear friend. Gracias, lindo, por todo el apoyo, comprensión, y amor que me mostraste durante toda mi educación universitaria. Los 15 años del camino de la vida que compartimos juntos fueron años llenos de muchos cambios, aventuras, desafíos, y aprendizajes—estos momentos forman parte integral de quién soy hoy día y han informado mi perspectiva sobre la lengua, el aprendizaje, y el multiculturalismo. Gracias, lindo.

I would also like to thank Karen Jordan, my grade school Spanish teacher who inspired me to keep trying to learn Spanish, despite my language difficulties in middle school with reading and writing. As it turned out, studying Spanish helped me to better develop my reading and writing skills in English too by providing me with a tool with which to more objectively understand and regard language as a structure of analysis. Ms. Jordan's passion for language teaching and unflagging support of my learning fueled the confidence in myself I needed to step out of my comfort zone to pursue my passion for travel and language learning. Gracias, maestra Karen.

I also must thank my two wonderful parents, Sharon and Jim Black. I feel that I am very much a fusion between my two parents. Although they have been separated for many years now, they have collectively contributed to my development as a person and as an educator. From my mother, a teacher herself, I have learned the patience, organizational skills, and grit that has helped me to persevere and finish my education. As a single parent raising two children while working full-time, attending night school, and even training for competitive cycling, my mother has been the biggest champion of my success and has continually been there to support me through my struggles and celebrate my achievements. Thank you, mom. From my father, I have acquired a passion for adventure and a level of open-mindedness that has continually served me to find new connections between ideas and seek out and value different ways of viewing the world. Thank you, dad.

Finally, I would like to thank my life partner, Cédric Albert Jacques Houy, for his support during these final two years of my education and for helping me to adapt and feel at home here in France. Being multilingual himself, he intimately understands the daily struggles of learning to live in a new language and culture and he has been so thoughtful and understanding of my language journey as I strive to adapt to French as my new language and France as my new home. We often find ourselves code-switching between French, Spanish, English, and sometimes even Italian as we are fortunate to share these languages in common. I look forward to our many future adventures together. Je t'aime, chéri.

I am so thankful to have had this opportunity to dedicate these years of my life to the passionate pursuit of knowledge from a variety of avenues—through my students and teaching experiences, through my professors and mentors, through my coursework, and now through my own research. I have been so fortunate to find my passion in this research and in this community. Language learning is at the center of what drives me to make a contribution to society and I endeavor to ensure that my future teaching and research career strives to successfully move more people along this path of meaningful life-long language learning.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1: Introduction

Adult language learners demonstrate extensive variation in performance and ultimate attainment that has been related both to task demands (Saeedi, 2020; Michel, Révész, Shi, & Li, 2019; Skehan & Shum, 2014) and to individual learner differences (Andringa & Dąbrowska, 2019; Dörnyei & Skehan, 2008). Late (i.e., adult) learners acquiring an additional language in the formal or instructed setting are believed to rely more extensively on explicit knowledge and declarative memory than those acquiring a new language in a naturalistic setting or at an earlier age (Paradis, 2004). Both the age of onset and the manner in which an individual acquires a new language have a direct impact on how target language input is processed, represented, and retrieved, implicating the involvement of distinct knowledge and memory systems (Paradis, 2004, 2009; Ullman, 2001, 2004, 2005, 2015). This study therefore aims to examine the impact that differentiated task demands (time constraint and stimuli presentation modality) and learner differences (motivation, attitudes about the Spanish language and target language community, metacognitive awareness, metalinguistic awareness of the target structure, tested and self-reported Spanish proficiency, average weekly Spanish use in and outside of the classroom, prior linguistic repertoire, and multilingualism) may have on performance with grammatical gender as a problematic feature for the late acquisition of morphosyntax (McCarthy, 2008; Montrul, Foote, and Perpiñán, 2008; Fernández-García, 1999).

1.1 Target structure of analysis: grammatical gender in Spanish

In the present study, grammatical gender was chosen as the target structure for analysis of linguistic performance for two principled reasons: 1. Grammatical gender is particularly problematic for adult learners of Spanish, leading to persistent and fossilized errors despite overall increasing proficiency (e.g., Fernández-García 1999, McCarthy 2008, Montrul et al., 2008); and 2. Grammatical gender presents several inherent features—namely morphological marking, noun gender class, and domain of agreement—with which to examine differentiated performance not only overall, but with specific components of the morphosyntactic structure in question as this may be relevant when considering how both native and nonnative language users may differentially process linguistic input (see: Bruhn de Garavito & White, 2002; Bruhn de Garavito, 2007).

When we examine grammatical gender as a typological feature of language, we find that it is present in some form in approximately twenty-five percent of world languages (Corbett, 1991). Grammatical gender is a specific form of a larger noun class system that allows for the formation of agreement between nouns and other linguistic structures such as adjectives, articles, pronouns and/or verbs (Corbett, 1991). World languages that feature grammatical gender include Spanish, French, Italian, Portuguese, Russian, German, Arabic, and Hindi, among others. Common systems of gender division include masculine/feminine, masculine/feminine/neuter, animate/inanimate, and common/neuter.

Grammatical gender in Spanish, like most Romance languages, is a binary system in which all nouns are assigned as masculine or feminine. Although gender assignation is a lexical property of nouns, grammatical gender is realized at the syntactic level in which there must be agreement between a noun and its determiner and modifier(s), thus resulting in two domains of grammatical gender in Spanish: assignment (lexical) and agreement (syntactic) (Alarcón 2009; 2011). The feature of animacy also plays a role at the lexical and semantic level in grammatical gender assignment. Animate nouns are those in which gender is assigned in accordance with biological sex and thus is semantically motivated (e.g. *el doctor* "the-masc. doctor-masc.", *la doctora* "the-fem. doctor-fem."), whereas inanimate nouns are those whose gender is not semantically motivated and is purely grammatical (e.g. *el libro* "the-masc. book", *la silla* "the-fem. chair"), making their classification arbitrary as compared to that of animate nouns (Montrul et al., 2008). Finally, noun morphology is also a relevant linguistic feature of grammatical gender in Spanish in which nouns can also be classified based on their level of morphological *transparency* for gender. Most nouns in Spanish follow a canonical or prototypical pattern in which their respective morphology reveals their grammatical gender class, such that masculine nouns tend to end in /-o/ and feminine nouns tend to end in /-a/ (e.g., *el libr**o*** "the-masc. book"; *la sill**a*** "the-fem. chair") (Montrul et al., 2008). This morphologically prototypical group of nouns can be described as canonical or overt, with regards to their morphology. There are other nouns that do not follow this prototypical pattern and therefore can be classified as non-canonical or non-overt because their morphology does not directly reveal information about their grammatical gender class.

These morphological variants include nouns that end in /-e/ (e.g., *el puente* “the-masc. bridge”) and in consonant (e.g., *la flor* “the-fem. flower”). There is also a subclass of non-overtly marked nouns which can be classified as exceptional in that, within this category, masculine nouns end in /-a/ (e.g., *el problem**a*** “the-masc. problem”) and feminine nouns end in /-o/ (e.g., *la man**o*** “the-fem. hand”), thus contradicting the prototypical pattern (Montrul et al., 2008; Alarcón, 2011; Foote, 2015). Nonetheless, according to Teschner and Russell (1984), 99.87% of all nouns that end in /-o/ are masculine and 96.30% of all nouns ending in /-a/ are feminine in the official *Diccionario de la Lengua Española*. Therefore, these exceptional cases are indeed rare and have therefore been excluded from the present study.

1.2 Aims and general research questions

This study is specifically motivated by a notable lack in previous research that examines how both native and nonnative language users respond to varied task demands. It takes a variationist approach that prioritizes variation and individual differences as key components for understanding the complex system of language development. Furthermore, this study endeavors to extend analysis beyond the prototypical monolingual-bilingual dichotomy by including and analyzing the performance of multilinguals and crucially considering multilingual experience as an independent explanatory variable for predicting performance with a particularly problematic structure. This more inclusive multilingual approach recognizes the role of individual linguistic experience and prior linguistic repertoire—as opposed to merely discussing ‘other languages reported’ as a limitation for analysis—by considering a potential ‘multilingual effect’.

Therefore, this doctoral thesis research examines how adult native speakers and learners of Spanish perform with grammatical gender during different tasks. Accuracy and variation in scores during performance are both analyzed and compared across different tasks designed to variably condition the use of explicit and implicit linguistic knowledge, as these factors may variably modulate performance and may be indicative of the distinct memory and knowledge systems employed by both native speakers and adult instructed language learners (N.C. Ellis, 2015; R. Ellis, 2005; Ullman, 2005; Paradis, 2004). A dynamic systems approach (Larsen-Freeman & Cameron, 2008) is adopted to measure performance in terms of both grammatical accuracy and variation, measured as inter-speaker group variation (SD) and intra-speaker variance, as two dependent variables that indicate the current state of development of the dynamic language system in both native and nonnative speakers (i.e., learners). The approach that this study takes recognizes the fundamental role that individual differences play in language performance and attainment and intentionally diverges from the idealized native speaker norm by analyzing performance and variation in both native and nonnative speakers, critically recognizing linguistic diversity in and among language users by analyzing the effect of multilingualism on novel language performance. As such, this study is guided by the following general research questions:

> **QUESTION 1.** How do native Spanish speakers and late learners of Spanish perform with grammatical gender agreement?
>
> **QUESTION 2.** How do the task demands (time constraint, input modality) impact accuracy and variation?
>
> **QUESTION 3.** To what extent can individual differences predict or account for accuracy and variation in performance?
>
> **QUESTION 4.** Is there a multilingual effect? Do participants who report experience with more languages perform differently from those who report experience with only English and Spanish?

This doctoral thesis helps elucidate how native and nonnative speakers' performance is impacted by linguistic processing demands imposed by constraints of task structure and modality. Furthermore, this research contributes to our understanding of how and to what extent adults acquire an additional language in the instructed setting and has important implications for language teaching methodology by uncovering the individual, linguistic, and task factors that most contribute to accuracy, variability/systematicity in performance, and more implicit nonnative language processing among learners from diverse linguistic backgrounds.

1.3 Outline

Chapter 2 – Theoretical framework and previous research provides an overview of the theory and previous empirical findings on which this study is based, organized by the different subdomains of research that this dissertation covers, including: language processing modality and differentiated memory systems (*section 2.1*), task demands and input modality (*section 2.2*), individual differences that impact performance and ultimate attainment (*section 2.3*), the variationist approach to language acquisition research (*section 2.4*), the connection between an emergentist/variationist and the generativist perspectives (*section 2.5*), and an examination of grammatical gender and its acquisition (*section 2.6*). In Section 2.7, a brief summary of previous research and theory as it pertains to each research question is presented along with the specific aims of the study based on these questions and motivated by the gaps left in previous literature. Finally, in Section 2.8, Chapter 2 concludes with a detailed presentation of the specific research questions that this study aims to address along with corresponding predictions grounded in previous research and theory.

Chapter 3 – Methodology presents an overview of the methods used to carry out this study. In Section 3.1, a full description of the study participants is provided, including the incidence of multilingualism observed in the sample (i.e., frequency of multilingualism and additional languages reported). In Section 3.2, the study procedure is described. Section 3.3 describes the study materials and tasks that participants completed, including the pre-experimental tasks (LOI + Consent, questionnaire), and presents the data collected on the Language Learner Profile Questionnaire, including: first language, language learning history, language use, self-reported language proficiency in all languages known, language attitudes, language learner awareness (metacognitive awareness), and motivational orientation. Participants' results on the Spanish proficiency test are also reported and each experimental task (Tasks 2-5) and the metalinguistic awareness exit survey are described in full, including illustrated examples of each task. Finally, in Section 3.4, the dependent/response variables and the independent/explanatory variables are presented and discussed, and both the descriptive and inferential data analysis procedures are explained.

Chapter 4 – Results: Linguistic and task effects presents the descriptive and inferential results according to the first two research questions on the effect of the linguistic and task variables. Section 4.1 details how both native Spanish speakers and adult learners of Spanish performed with grammatical gender on all tasks and presents the results according to the effect of the linguistic variables per speaker group (native speaker and adult/late learner) and per proficiency group (beginner, intermediate, advanced). Section 4.2 presents the results of task effects in both speaker groups and per learner proficiency level, including the effect of a time constraint (i.e., speeded tasks vs. self-paced tasks) and input modality (auditory stimuli vs. written stimuli). Finally, Chapter 4 concludes with a brief descriptive summary of the key findings uncovered regarding the effects of the linguistic and task variables in both native Spanish speakers and adult learners.

Chapter 5 – Results: Individual factors and multilingual effect presents the descriptive and inferential results according to the second two research questions on the effect of the individual factors and multilingualism. In Section 5.1, results are presented according to the individual factors analyzed, treated both categorically, in which participants were divided into separate groups for each factor, as well as continuously, based on scores on each factor, including: Spanish proficiency, metacognitive awareness, motivational orientation, metalinguistic awareness, attitudes (about the Spanish language and target language community), Spanish use, and prior linguistic repertoire (presence of grammatical gender in other reported languages). A multiple linear regression model is then presented for each speaker group demonstrating the combined effect on performance of the individual factors analyzed. Section 5.1 concludes with a presentation of a Pearson's *r* correlation matrix demonstrating how the individual factors analyzed are correlated with one another in order to examine more closely the nature of individual differences in adult learners. Section 5.2 details the results of the multilingual effect in both participant groups in which accuracy scores are compared between multilinguals and bilinguals both overall and on each task type. The multilingual effect is further explored through an analysis of the effect of typology of the additional language(s) reported, categorized according to the presence of grammatical gender, and furthermore, binary grammatical gender. The results of how the number of additional languages reported affects performance are presented for both native speaker and late learner groups. Finally, Chapter 5 concludes with a brief descriptive summary of the key findings uncovered regarding the relative effect of each individual factor on performance and the multilingual learner advantage.

Chapter 6 – Discussion first summarizes and discusses in Section 6.1 the key results in light of the initial predictions made for each research question. In Sections 6.2, 6.3, 6.4, and 6.5 the results for native and nonnative performance with gender agreement and the relative effect of each linguistic variable, task variable, individual factor, and multilingualism, respectively, are further discussed in relation to the predictions made, and possible explanations are proposed based on previous research and theory. Section 6.6 presents the output of comprehensive multiple linear regression modeling that integrates all independent variables analyzed to predict performance (accuracy) with grammatical gender: four models for nonnative Spanish learners as a group and per proficiency level and one model for native Spanish speakers. Two key questions are then discussed: What do these models tell us? What can these models not tell us? A brief summary of the key findings and how they can be supported or explained by theory is presented in Section 6.7. Then, in Section 6.8, we examine to what extent the findings of the present study can answer the *learnability* question of if grammatical gender is acquirable by adult learners and further consider what factors uncovered in this study may be able to explain this (varied) outcome. Section 6.9 details and discusses the limitations of the present study and, finally, Section 6.10 considers possible avenues of future research based on the findings and limitations of this study.

Chapter 7 – Conclusion revisits the aims of the present study and briefly summarizes the key findings for all research questions taken together. The final chapter of this dissertation concludes with a discussion of the significance and contributions of this study to the larger fields of adult language acquisition and multilingualism and considers the pedagogical implications of these findings.

Finally, the references are provided followed by a complete Appendix including the study materials (Language Learner Profile Questionnaire, Spanish proficiency test, experimental tasks) and individual tables presenting the complete descriptive data for each task participants completed. In the next chapter, previous research and theory will be presented and briefly discussed to contextualize and justify the present study.

Chapter 2: Theoretical framework and previous research

2.1 Language processing modality and differentiated memory systems

Implicit-explicit language knowledge

There are two main modes of language learning that both occur to some degree in all learners and whose relative proportions in any given learner depend principally on age and context. *Implicit* language learning refers to the acquisition of underlying structure from a complex stimulus environment by a process which takes place naturally, simply, and without conscious awareness. *Explicit* language learning, on the other hand, constitutes a more conscious operation in which the individual makes and tests hypotheses in an overt search for structure (N.C. Ellis, 2015). Therefore, acquiring language competence can take place implicitly via the nonconscious and automatic abstraction of structural patterns from experiencing language (i.e., meaningful input) or explicitly through selective learning in which the learner searches for information and tests hypotheses also based on language input (N.C. Ellis, 2015). According to N.C. Ellis (1996), implicit learning is primary for linguistic competence, whereas explicit learning of language is typically the end product of acquisition—not its cause. Implicit language learning is our natural default setting for processing and integrating linguistic input; when we use language, we are conscious of communicating rather than of counting or computing, yet in the course of conversation we also naturally (and implicitly) acquire knowledge of the frequencies of different linguistic elements, their relative dependencies, and their mappings as part of a larger, underlying systematic structure such that linguistic knowledge necessarily involves underlying statistical knowledge (Bybee & Hopper, 2001; N.C. Ellis, 2002; Gries & Divjak, 2012).

R. Ellis (2005) distinguishes implicit and explicit linguistic knowledge across a variety of parameters including awareness, knowledge type, degree of systematicity, nature of accessibility and use, ability to self-report, and learnability. From the perspective of these distinct parameters, *implicit knowledge* involves *intuitive* awareness of linguistic norms, *procedural*-type knowledge of rules and fragments, is more *systematic*, accessed by means of *automatic* processing during *fluent* language use, is *nonverbalizable*, and is most readily and consistently learnable within a *critical or sensitive period*. In contrast, *explicit knowledge* involves *conscious* awareness of linguistic norms, *declarative*-type knowledge of grammatical rules and fragments, is *anomalous* and *variable*, accessed by means of *controlled* processing requiring some degree of *planning*, is *verbalizable*, and is *learnable at any age* (R. Ellis, 2005). A summary of the different characteristics of implicit and explicit knowledge as explained by R. Ellis (2005) is presented in Table 1 below. Both generativist and cognitivist accounts of language acquisition acknowledge that *linguistic competence*—and not knowledge about language—crucially comprises *implicit knowledge* such that language acquisition is evident in what language users know *intuitively* (Gregg, 1989; Krashen, 1985).

Table 1. Key characteristics of implicit and explicit knowledge according to R. Ellis (2005).

Characteristics	Implicit knowledge	Explicit knowledge
Awareness	*Intuitive* awareness of linguistic norms	*Conscious* awareness of linguistic norms
Type of knowledge	*Procedural* knowledge of rules and fragments	*Declarative* knowledge of grammatical rules and fragments
Systematicity	Variable but *systematic* knowledge	Anomalous and *inconsistent* knowledge
Accessibility	Access to knowledge by means of *automatic processing*	Access to knowledge by means of *controlled processing*
Use of L2 knowledge	Access to knowledge during *fluent* performance	Access to knowledge during *planning* difficulty
Self-report	*Nonverbalizable*	*Verbalizable*
Learnability	Potentially only within *critical period*	*Any age*

Adapted from R. Ellis (2005) "Key characteristics of implicit and explicit knowledge".

Although implicit learning is the norm in native (L1) acquisition, the overwhelming consensus in late acquisition research is that nonnative adult acquisition by implicit means alone is limited in its success (e.g., Schmidt, 1990; Lightbown, Spada, & White, 1993). In other words, much of the *input*—the available target language in the environment—does not necessarily become *intake*—the subset of input which is processed and integrated into the developing linguistic system (see: "Input Processing Theory" and "Processing Instruction", VanPatten, 1996, 2015; Corder, 1967). The L2 learning literature abounds with many instances of how years of input can fail to become intake (e.g., Perdue, 1993; Klein, 1998) and suggests that implicit learning mechanisms (e.g., statistical tallying) do not take place in cases where linguistic form lacks perceptual salience (Schmidt, 1990, 2001) or where the L2 semantic/pragmatic concepts to be mapped onto the L2 forms are unfamiliar. In these cases, additional attention is necessary for the relevant associations to be learned and this can be achieved through form-focused instruction that recruits learners' explicit conscious processing (N.C. Ellis, 2005). Years of learned attention during native language and general cognitive development limit the potential of implicit learning for subsequent language acquisition. Therefore, explicit learning, and at least some explicit instruction, seems to be necessary to reach target-like norms in late nonnative language development (N.C. Ellis, 2002). However, as N.C. Ellis & Larsen-Freeman (2009) argue, the complex adaptive system of interactions within and across form and function is far richer than that emergent from implicit or explicit learning alone as interactions of conscious and unconscious learning processes play roles at all emergent levels. Language is, therefore, a complex, dynamic and adaptive system in which interactions between the implicit and the explicit are constant currents (Beckner et al., 2009; N.C. Ellis, 2007; N.C. Ellis & Larsen-Freeman, 2006, 2009; Larsen-Freeman, 1997).

There has been debate on whether and to what extent implicit and explicit knowledge systems may be related (see: Bialystok, 1994; R. Ellis, 1993, 2005; Krashen, 1981; Paradis, 1994; Hulstijn, 2002) and three different theoretical perspectives exist: 1. the *non-interface position*; 2. the *strong interface position*; and 3. the *weak interface position*. The *non-interface position* posits that implicit and explicit L2 knowledge involve different acquisitional mechanisms (Hulstijn, 2002; Krashen, 1981), are stored in different parts of the brain (Paradis, 1994), and are accessed for performance by different processes, either automatic or controlled (R. Ellis, 1993). This position rejects both the possibility of explicit knowledge transforming directly into implicit knowledge and the possibility of implicit knowledge becoming explicit. However, a weaker form of the non-interface position recognizes the possibility of implicit knowledge transforming into explicit knowledge through the process of conscious reflection on and an analysis of output generated by means of implicit knowledge (Bialystok, 1994).

Both the strong interface position and the weak interface position, in contrast, recognize a connection between the two knowledge systems. The *strong interface position* makes two key points: 1. that explicit knowledge can be derived from implicit knowledge, and 2. that explicit knowledge can be converted into implicit knowledge through practice (M.S. Smith, 1981; DeKeyser, 1998). In other words, learners can first learn the rules as a declarative fact and then, through practice, convert this same knowledge into an implicit representation. The *weak interface position* exists in three versions, all of which acknowledge the possibility of explicit knowledge becoming implicit but posit some limitations on when or how this can take place. The first version argues that explicit knowledge can convert into implicit knowledge through practice only if the learner is developmentally ready to acquire the linguistic form (R. Ellis, 1993). The second version of the weak interface position argues that explicit knowledge contributes only *indirectly* to the acquisition of implicit knowledge by promoting some of the processes believed to be responsible. For example, N.C. Ellis (1994) suggested that “declarative rules can have ‘top-down’ influences on perception” (p. 16), in particular by making relevant features salient and thus enabling learners to notice them and to notice the gap between the input and their existing linguistic competence. Finally, the third version of the weak interface position argues that learners can use their explicit knowledge to produce output that then serves as auto-input to their implicit learning mechanisms (Schmidt & Frota, 1986; M.S. Smith, 1981).

Independently of which (non-)interface position you adopt, there is wide acceptance that explicit knowledge can contribute to performance. For example, Krashen (1977) argued that explicit knowledge is available to the *monitor*—the production mechanism that enables learners to edit their own performance by drawing on what they consciously know to be correct. Furthermore, Bialystok (1982) showed that different performance tasks are likely to induce L2 learners to draw differentially on their implicit and explicit knowledge. For example, formal writing tasks are likely to induce learners to draw more extensively on their analyzed (explicit/declarative) knowledge of an L2 than tasks that solicit unplanned, oral communication. It is with this association in mind between task type and knowledge type that the present study has been designed to examine differentiated task demands as a potential explanatory variable for variation in adult/late learner performance.

In summary, different iterations of the implicit-explicit interface debate recognize that all language users possess both implicit and explicit linguistic knowledge, albeit in different proportions, and that both knowledge types have the potential to develop in tandem and simultaneously, although age, context, and manner of acquisition necessarily influence their relative proportions (M.S. Smith, 1981; DeKeyser, 1998; N.C. Ellis, 1994, Hulstijn, 2002).

Frequency effects in nonnative language acquisition

Necessarily related to the vein of implicit and explicit language knowledge and processing is the role of frequency in language acquisition. N.C. Ellis (2002) argues that because the conscious (i.e., explicit) experiences of language learning do not revolve around counting, to the extent that language processing is based on *frequency* and *probabilistic* knowledge, language learning is *implicit* learning. Language acquisition, representation, and processing are all tuned to varying degrees by frequencies in the input (N.C. Ellis, 2002). Previous research has detected significant positive correlations between accuracy and frequency (e.g., Brown, 1973), which has subsequently been corroborated by ESL teacher talk data during classroom instruction (Larsen-Freeman, 1978). Other researchers have uncovered correlations between the frequency of occurrence in learner-directed input and developmental sequences in SLA (Hatch & Wagner-Gough, 1976). Frequency effects can be examined in two main domains: *type frequency* and *token frequency* (N.C. Ellis, 2002). Token frequency is how often particular words or specific phrases appear in the input, whereas type frequency is how many different lexical items can be applied to a certain pattern, paradigm, or construction.

Ellis and Schmidt (1997) investigate adult acquisition of L2 morphology using an artificial language in which frequency and regularity were factorially combined. Participants first learned 20 new names for pictures and then learned the plural forms for these names. Half of the items were regular plurals in that they shared the same plural affix and the other items exhibited unique plural affixes. In addition, half of the regular and irregular forms were high frequency in that they were presented five times more than the other items. On each trial, the correctness and reaction time (RT) of the learner’s verbal response were assessed. The results indicated an advantage for high-frequency items and this frequency effect was larger for the irregular items. In other words, when a structure is less systematic (i.e., contradicts the canonical pattern), frequency effects become more pronounced. Ellis and Schmidt (1997) uncovered positive frequency effects on both accuracy and RTs for regular/canonical items in the early stages of acquisition. Furthermore, the effect size of frequency was found to diminish with learning (i.e., increased overall proficiency). Finally, the size of the frequency effect on irregular items also diminished with increasing proficiency, but at a slower rate. In sum, this study uncovered frequency-by-regularity interactions; when a grammatical item is irregular, frequency effects are more pronounced.

Frequency is apparently important as it has been found to be highly correlated with accuracy, but the question remains of *why*? Frequency is likely a placeholder, or rather a correlate, of something more profound occurring in the mental grammar of the individual language learner. An interpretation of the effect of frequency depends on one's conception of the language acquisition process (Larsen-Freeman, 2002). One possible explanation put forth by Larsen-Freeman (2002) is that structures with high frequency might be more *noticeable* and therefore supply learners with more data on which to base and refine hypotheses about rules. In connectionist models of neural networks, items that are frequent in the input increase the connection weights between nodes, which are taken to represent neurons in the central nervous system. As the input is processed, certain connection weights are strengthened, while others atrophy, simulating the plasticity of synaptic connections. Larsen-Freeman (2002) emphasizes that simple frequency does not in and of itself explain much about language acquisition; the role of frequency needs interpreting. In other words, a frequency effect cannot be interpreted by the numbers alone. Recognizing this, N.C. Ellis (2002) qualifies the impact of frequency in three ways: 1. frequency does not always yield a uniform effect; 2. there is a moderating influence of cue competition and constraint satisfaction; and 3. there is a difference between type frequency and token frequency.

Nonetheless, important limitations of frequency effects have been uncovered. Larsen-Freeman (2002) notes that "[…] although we are noticers of probabilistic tendencies, we are also sense-makers" (p. 6). In other words, frequency can only have an effect if the input is sufficiently comprehensible. In this sense, it seems that it would interact with overall developing proficiency such that frequency effects are strongest at somewhat intermediate levels of proficiency; the learner needs to understand enough of the input in order to categorize and tally comprehended structures for implicit analysis and subsequent integration into their developing linguistic system. In addition, formal instruction can also modulate natural frequency effects in grammar learning. Relative frequency of certain grammatical items in instructional input can differ from natural frequency as teaching may focus on certain forms deemed problematic/difficult precisely because they are less frequent in natural input. Furthermore, Gass and Mackey (2002) highlight how frequency effects can be dynamic throughout the acquisition process as they are modulated to varying degrees by proficiency via the importance of "noticing"; initially, noticing patterns is most important with frequent items and then gradually becomes more important with infrequent items as acquisition progresses. Gass and Mackey (2002) further relate maturational constraints to frequency effects. They argue that it is possible that the age effects observed in much empirical research actually relate to sensitivity to frequency of input (in implicit learning) and that this may also help to explain differences between L1 and L2 acquisition.

Now that we have considered both implicit and explicit language knowledge and processing, including frequency effects in implicit linguistic competence, we will consider next two different memory systems involved in nonnative language acquisition and learning.

Declarative-procedural model

Cognitive and neurological research has posited roles for two long-term memory systems—*declarative* and *procedural*—in both first and second language learning (Morgan-Short & Ullman, 2012; Ullman, 2005, 2015). Declarative and procedural memory systems differ along a number of dimensions, including the role of awareness, the computations they perform, and the neural substrates that subserve them (Eichenbaum, 2002; Eichenbaum & Cohen, 2001). In general cognitive terms, declarative memory supports the learning of general facts and knowledge whereas procedural memory supports motor and sequential skills (Knowlton & Moody, 2008). Furthermore, procedural memory consists of implicit knowledge since the knowledge contained therein is difficult to verbalize and access via introspection. In terms of language functions, declarative memory is involved in the acquisition of the mental lexicon whereas procedural memory is involved in the acquisition of the mental grammar (Paradis, 2004, 2009; Ullman, 2001, 2004, 2005, 2015). Ullman's declarative-procedural (DP) model formalizes this distinction for first language development. According to the DP model, declarative memory underlies the acquisition and representation of information stored in the lexicon, including words and grammatically complex forms that are memorized as whole chunks (due to their frequency). In contrast, procedural memory is posited to underlie aspects of grammar thought to rely on combinatorial processing, such as morphosyntax and syntax (Ullman, 2004, 2005).

In L2 development, the DP dichotomy is hypothesized to be somewhat different. As in L1, L2 lexical development is argued to rely on declarative memory; however, in contrast to L1 grammar, early L2 grammar development is argued to rely on declarative memory as well (Hamrick, 2015), and only at higher proficiency levels can grammar learning take place in the procedural system (Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter, & Wong, 2014). Increasing evidence from electrophysiology (e.g., Morgan-Short, Finger, Grey, & Ullman, 2012; Morgan-Short, Steinhauer, Sanz, & Ullman, 2012) and neuroimaging (Morgan-Short, Deng, Brill-Schuetz, Faretta-Stutenberg, Wong, P.C.M., & Wong, F.C.K., 2015; Tagarelli, 2014) supports these predictions and individual differences in declarative and procedural memory abilities have been shown to correlate with L2 learning (Carpenter, 2008; Morgan-Short et al., 2015). More specifically, Morgan-Short et al. (2014) provide empirical evidence for the declarative-procedural memory distinction in L2 learning as they found that behavioral measures of declarative memory predicted grammar learning early in language training, but not in later stages of training. In contrast, measures of procedural memory predicted grammar abilities in later, but not earlier, phases of language training and these findings are consistent with the theoretical predictions of the DP Model.

Online-offline language processing

Important information about the relative balance between declarative and procedural memory systems in nonnative language performance can be gleaned from how learners are conditioned to respond to particular linguistic stimuli. Cognitively *online* and *offline* language processing can be variably conditioned in language tasks so as to differentially draw more heavily on either declarative or procedural memory systems. Cognitively *online* tasks are those which are more spontaneous, more complex, and that do not allow for planning time (Skehan & Foster, 1997). Online language processing requires the use of *implicit* knowledge stored in the *procedural* memory system, while *offline* processing allows the language learner time to tap into their *explicit* language knowledge stored in the *declarative* memory system. This online-offline distinction in language processing arises from the fact that second language users have limited attentional capacities such that task structure, complexity, and planning time variably influence accuracy rates in performance. With regards to the learning of grammatical gender, the structure under analysis in the present study, Prévost and White (2000) claim that L2 learners have the feature *gender* represented in their L2 grammars at an abstract level, but that nonetheless gender errors still occur due to an assembly or production problem or a computational difficulty during online processing (i.e., during unplanned production), as argued in their Missing Surface Inflection Hypothesis. Subsequent empirical research seems to corroborate this hypothesis with higher accuracy rates observed during cognitively offline structured written tasks than during online unplanned oral production tasks (Gamboa, 2012; Grüter, Lew-Williams, & Fernald, 2012; Alarcón, 2011; Sagarra & Herschensohn, 2010; Montrul et al., 2008). It is with these insights in mind that the present study aims to examine language performance across varied tasks that have been designed to condition online and offline language processing.

Issues of measurement: operationalizing the implicit-explicit distinction

Although a theoretical distinction between implicit and explicit linguistic knowledge clearly exists in the literature and neuroimaging studies have uncovered a declarative-procedural dichotomy in long-term memory, operationalizing this distinction for SLA research has proven technically difficult. R. Ellis (2005) proposes seven criteria by which to operationalize the implicit-explicit distinction for empirical research: degree of awareness, time available, focus of attention, systematicity, certainty, metalinguistic knowledge, and learnability. Table 2 summarizes the operationalization of these criteria to solicit and evaluate implicit and explicit L2 knowledge according to R. Ellis (2005).

Table 2. Operationalizing the constructs of L2 implicit and explicit knowledge.

Criterion	Implicit Knowledge	Explicit Knowledge
Degree of awareness	Response according to feel	Response using rules
Time available	Time pressure	No time pressure
Focus of attention	Primary focus on meaning	Primary focus on form
Systematicity	Consistent responses	Variable responses
Certainty	High degree of certainty in responses	Low degree of certainty in responses
Metalinguistic knowledge	Metalinguistic knowledge not required	Metalinguistic knowledge encouraged
Learnability *(the extent to which knowledge can be internalized by learners at a given stage of development)*	Early learning favored	Late, form-focused instruction favored

Adapted from R. Ellis (2005) "Measuring implicit and explicit knowledge of a second language: A psychometric study".

Implicit knowledge is encouraged or conditioned in tasks which solicit a response according to feel (i.e., intuition, when something just sounds right), involve a time pressure, are focused on meaning, produce consistent responses with little intra-speaker variation, usually lead to a higher degree of certainty in responses, do not require metalinguistic knowledge to successfully perform, and favor early language learning. In contrast, explicit knowledge is encouraged or conditioned in tasks which solicit a response according to rules, are self-paced (i.e., no time pressure), are primarily focused on form, produce variable responses with a lower degree of certainty, encourage the use of metalinguistic knowledge, and favor form-focused instruction (see Table 2).

In line with these criteria, R. Ellis (2005) developed a battery of tests consisting of those which are intended to solicit implicit knowledge—including an oral imitation task, an oral narrative task, and a timed grammaticality judgement task (GJT)—and explicit knowledge—including an untimed GJT and a metalinguistic knowledge test. A summary of the design features used to develop R. Ellis' (2005) battery of tests to assess implicit and explicit linguistic knowledge is provided in Table 3 below.

Table 3. Design features of the battery of tests developed by R. Ellis (2005).

Criterion	Imitation Task	Oral Narrative Task	Timed GJT	Untimed GJT	Metalinguistic Task
Degree of awareness	Feel	Feel	Feel	Rule	Rule
Time available	Pressured	Pressured	Pressured	Unpressured	Unpressured
Focus of attention	Meaning	Meaning	Form	Form	Form
Metalinguistic knowledge	No	No	No	Yes	Yes

Adapted from R. Ellis (2005) "Measuring implicit and explicit knowledge of a second language: A psychometric study".

The key design features of these tasks include degree of awareness, time available, focus of attention, and metalinguistic knowledge. The implicit knowledge tests (i.e. imitation, oral narrative, and timed GJT) solicit a response according to feel or intuition, are administered under a time pressure, focus attention on meaning, and do not require metalinguistic knowledge; in contrast, the explicit knowledge tests (untimed GJT and metalinguistic knowledge test) solicit a response based on a rule, are self-paced (i.e. no time pressure), focus attention on linguistic form, and require some metalinguistic knowledge to successfully complete. The tasks for the present study have been selected and adapted from R. Ellis' (2005) battery of tests and are described in detail in the Methodology section.

Nonetheless, it is also important to discuss the inherent limitations in operationalizing this distinction between tests of implicit and explicit knowledge. First and foremost, there can be no guarantee that the "task-as-workplan" will directly and consistently correspond to the "task-as-process", as argued by some researchers (e.g., Breen, 1989; Coughlan & Duff, 1994). In other words, learners are likely to draw on whatever resources they have at their disposal irrespective of which resources are the ones best suited to the task at hand. At best, then, tests designed to operationalize this distinction are expected to *predispose* learners to access one or the other knowledge type, but only in a *probabilistic* manner. This important limitation will be reconsidered in the Discussion section concerning possible task effects and how this may be indicative of different knowledge and memory systems employed by late learners.

2.2 Task demands and input modality

Task condition effects: demands & task type

Different performance tasks are likely to induce L2 learners to draw differentially on their implicit and explicit knowledge (Bialystok, 1982). For example, formal writing tasks are likely to induce learners to draw more extensively on their analyzed explicit knowledge than tasks calling for unplanned, oral communication. Research has consistently shown a significant relationship between task variables such as structure/organization, familiarity, personal relevance, stimuli presentation modality, and planning time and different measures of linguistic performance such as accuracy, complexity and fluency (e.g. Wang & Skehan, 2014; Skehan & Shum, 2014; Tavakoli & Foster, 2008). Task structure that lowers the demands on the conceptualization stage of speech production (see Levelt's 1989 speech production model) frees up attentional resources for allocating to formulation processes, thereby increasing accuracy in speech performance (Wang & Skehan, 2014; Skehan & Shum, 2014). However, in order to separate the effects of task-related factors from effects that are due to incomplete (Montrul, 2016) or different (Kupisch & Rothman, 2018) acquisition of the target language, it is essential to examine how task manipulations influence the linguistic performance of both native and nonnative speakers—something that few studies have endeavored to do (e.g., Michel et al., 2019; Foster & Tavakoli, 2009).

One of the few studies that compares native speakers to adult learners with regards to task effects is Michel et al. (2019). In their study they investigated the extent to which increasing cognitive task demands affect the syntactic complexity and accuracy of language performance across different speaker groups (L1 and L2) and task types. English L1 speakers (n=16) were compared with German L2 learners of English (n=16) in terms of their syntactic complexity and accuracy across cognitively less and more demanding versions of three task types: a decision-making task, a map task, and a narrative task. Michel et al. (2019) found that the effects of cognitive task demands on syntactic complexity and accuracy varied according to task type and speaker status such that L2 users produced more subordinate clauses with increased cognitive demands, yet for the native speaker group, the task demands had no effect on the decision-making task. Overall, varied effects were observed depending on task type such that no effect was detected for either speaker group on the map task, yet differences were detected between groups on the narrative task and the decision-making task.

Building off of the limited previous research in task effects per speaker/learner groups, the present study examines differentiated task demands as a potential explanatory variable for variation in both native speaker and learner performance as this may be indicative of the different memory systems and the nature of linguistic knowledge that native speakers and adult instructed learners draw upon when performing diverse language tasks.

Input/stimulus modality during testing: auditory vs. written

One important component that characterizes the demands of any given task is how the task stimuli are presented: auditory (listening) or written (reading). Research in cognitive science has provided evidence that the modality in which information is encoded affects how we process and learn the information (e.g., Crowder, 1986; Engle & Mobley, 1976). Aural (auditory) and visual input are processed differently, in different cortical areas, and the aural cortex displays more sensitivity to pattern detection whereas the visual cortex is superior for detecting spatial information (e.g., Chen & Vroomen, 2013; Frost, Armstrong, Siegelman, & Christiansen, 2015; Recanzone, 2009). Psychological research on modality differences suggests that verbal material presented aurally and visually is processed in different parts of the memory system and by different mechanisms (Penney, 1989). Furthermore, language studies suggest that the written modality enables more elaborate (e.g., Vasylets, Gilabert, & Manchon, 2017) and more accurate (e.g., Kormos, 2014) language production than the aural modality and is also perceived as less difficult by literate adult learners (e.g., Cho, 2018). Written input may free up attentional resources to notice, attend to, and potentially become more aware of target features in input. These claims have been corroborated by empirical language acquisition research that has shown higher accuracy rates (Bialystok, 1979, 1982; Johnson, 1992) and faster reaction times (Murphy, 1997) for written grammaticality judgement tasks (GJT) when compared to results obtained from GJTs presented in the aural modality (e.g., Haig, 1991; Johnson, 1992). Murphy (1997) explains this dichotomy in terms of the perceived "burdens of auditory processing" (Murphy, 1997, p. 55) and the ease of visual processing due to heightened cognitive demands imposed on language users when processing auditory stimuli as opposed to permanent visual stimuli. Overall, the findings from SLA research and cognitive science suggest that the presentation rate, that is, untimed for written and timed for aural, as well as the physical medium, visual or aural mode, may both differentially impact the L2 learning process and performance. Furthermore, findings from the SLA literature suggest that the untimed nature of written input allows for better information uptake than auditory input does (Bialystok, 1997, 1982; Johnson, 1992; Murphy, 1997).

Speaker status as native or nonnative also variably interacts with the effect of stimulus modality as detected in Murphy's (1997) study examining the effect of aural and visual modalities among L2 English and French learners when compared to their respective native speaker counterparts; although differences per modality were observed for both native and nonnative speaker groups, auditory processing seemed to produce a greater obstacle for L2 learners to overcome than for native speakers.

Furthermore, the nature of language instruction also plays a role in how learners perform with the aural and written modality types. Sydorenko (2010), for example, investigated the effect of input modality (video, audio, and captions) on the learning of written and aural word forms in beginner learners of Russian (N=26) and found that learners' performance was directly linked to the modality type of the instruction they received; the group with only the written captions scored higher on written than on aural recognition of new words while the group with audio only (no captions) scored lower on written than on aural recognition of new words, while the learner group who received instruction with both the audio and the written stimuli (video audio with captions) learned the most new words of all three instructional type groups. On the follow-up questionnaire, the learners in the study expressed that they paid the most attention to captions, followed by video and audio, and acquired the most words by associating them with their corresponding visual image (i.e., caption) of the same word. Therefore, Sydorenko (2010) concluded that L2 learners prioritize written input, but when written input is not provided, they pay attention to audio input and subsequently perform better with audio input than written input at testing. Therefore, the modality of input in training/instruction affects the preferred modality at testing and performance. Although the natural default preference for L2 learners appears to be the written modality, this preference may be malleable via instruction.

With regards to grammatical gender as a problematic feature in the late nonnative acquisition of morphosyntax investigated in this study, adult instructed learners tend to be more accurate with gender agreement in orthographic visual stimuli when compared to pictorial stimuli. This preference is partially due to increased cognitive demands and vocabulary constraints inherent in picture-only task prompts, but also likely in part due to sensitivities to morphological cues present in orthographic input—an effect particularly pronounced when nouns follow the prototypical pattern for gender marking in Spanish (e.g., Black & Tararova, 2020).

Time constraint: speeded vs. self-paced tasks & planning

Similar to the effect of stimuli presentation modality, language learners' performance can also be affected by how much time they are permitted to complete a given task. Information Processing Theory claims that humans possess a limited processing capacity and, as a result, are not able to attend fully to all aspects of a task simultaneously (Anderson, 1995; Newell & Simon, 1972). Language learners, especially those with limited proficiency, find it difficult to attend to meaning and form at the same time and thus have to make decisions about how to allocate their attentional resources by prioritizing one aspect of language over others (Anderson, 1995; Skehan, 1996; VanPatten, 1990). However, when language learners have the opportunity to plan linguistic and propositional content before and/or during language performance, they can compensate for these processing limitations and, as a result, the quality of their linguistic output is enhanced (Skehan, 1996). In this sense, planning helps learners to access linguistic material from memory more easily and rapidly, particularly items stored in declarative memory that by nature require greater working memory to retrieve.

Levelt (1989) proposed the L1 Speech Production Model, subsequently applied to L2 learning (Kormos, 2006, 2011), outlining four incremental and parallel stages of speech production: 1. the conceptualization stage, which entails the planning of what one intends to say, resulting in a pre-verbal message; 2. the formulation stage, during which the pre-verbal message is transformed into its corresponding linguistic form through the process of lexical retrieval and syntactic and phonological encoding; 3. the articulation stage, in which the phonemic representation activates phonological forms and retrieves articulatory gestures to prepare the actual speech utterance; and 4. the monitoring stage, which occurs at each of the previous stages operating in multiple feedback loops. While L1 speakers' formulation and articulation operations tend to be automatic, and thus can work in parallel with conceptualization processes, L2 learners' formulation will likely be less automatic and therefore take up greater attentional resources, given that L2 users tend to have a less developed mental lexicon and less advanced grammatical encoding skills (Kormos, 2006). In turn, competition arises for cognitive resources in L2 user speech, yielding trade-offs in the amount of attention L2 users can dedicate to conceptualization versus formulation processes, which might become observable as trade-offs between complexity and accuracy measures (Skehan, 2009). Later, Michel et al. (2019) proposed the Limited Attention Capacity Model positing that working memory capacity and attentional resources are limited, and therefore there will be a competition for attentional resources during speech production processes. Overall, task factors can make separate demands on conceptualization and formulation (Skehan, 2009, 2015) and, therefore, a key practical issue is how increased pressure on the conceptualizer and/or formulator will be reflected in performance, expressed in terms of the linguistic areas of complexity, accuracy, and fluency (Skehan, 2015).

Planning allows learners to compensate for increased cognitive demands, enhancing both the accuracy and complexity of their output. Planning can occur before (i.e., *strategic planning*) or during a task, or both. Planning that occurs during a task is considered *online planning* in which speakers carefully attend to the formulation stage and engage in pre-production and post-production monitoring of their speech acts (Michel et al., 2019). When there is limited or constrained time for online planning, *rapid planning* can still occur, involving a greater degree of improvisation (Ochs, 1979). Therefore, time is a central factor that influences performance due to the degree of planning that can occur and the extent to which online and rapid planning occur is continuous rather than dichotomous, depending on the amount of time available as well as other learner and structural factors. Yuan and Ellis (2003) found that among adult L2 learners, online planning enhanced complexity and accuracy in performance whereas strategic/pre-task planning enhanced fluency. Similarly, Wang (2014) found that a combination of both online and pre-task planning increased accuracy in learner speech performance.

However, empirical findings have also shown that planning during linguistic performance does have its limitations. Diminishing returns have been detected such that after a certain amount of planning time (e.g., five minutes), any additional time does not actually result in enhanced performance (Mehnert, 1998) and certain linguistic structures (Ortega, 1999) are more susceptible to the effects of planning than others. Furthermore, the nature of the examined structure—rule-based vs. item-based (e.g., irregular past tense forms in English)—can also influence the efficacy of planning time such that performance with rule-based structures tends to benefit more from planning time. This could possibly be due to learners' ability to access explicit knowledge of the rule in declarative memory when provided with sufficient time to do so (R. Ellis, 1987).

With these factors and limitations in mind, three of the experimental tasks in this study (speeded imitation task, speeded auditory GJT, and speeded written GJT) have been designed to create a time pressure that, by nature, will only allow for minimal online planning, while the other two experimental tasks (self-paced written GJT and metalinguistic exit survey) are self-paced to allow for a greater degree of careful and controlled speech planning and monitoring.

2.3 Individual differences that impact performance and ultimate attainment

Overview: factors affecting ultimate attainment in adult language learners

Adult SLA is characterized by large and varied individual differences and adult learners rarely attain target-like competence (Andringa & Dąbrowska, 2019). Critical differences between native and nonnative language acquisition have traditionally been attributed to a biologically determined critical or sensitive period and a large body of research has provided evidence for its existence (e.g., Abrahamsson & Hyltenstam, 2009; DeKeyser, 2012; Granena & Long, 2013; Hyltenstam & Abrahamsson, 2003). Nonetheless, patterns seeming to support the existence of a critical period could also be explained to some degree by a range of other factors, such as the quality and quantity of input, changing motivations and attitudes, and contextual factors (Birdsong, 2005, 2006; Birdsong & Vanhove, 2016; Singleton & Muñoz, 2011). Another way of looking at the age effect issue is that past the hypothesized critical period for language learning, age is no longer such a determining factor and other individual learner variables take on greater importance in predicting performance and ultimate attainment. The present study therefore endeavors to investigate the role of a diverse set of individual learner variables that have been found to predict to some extent performance in adult learners, particularly in the instructed context: motivation, attitudes, metacognitive awareness, metalinguistic awareness, target language use, prior linguistic repertoire, and multilingualism.

Motivation

Motivation concerns the direction and magnitude of human behavior, including the choice of a particular action, the persistence with it, and the effort expended on it. In other words, motivation is responsible for why a person decides to do something, how long they are willing to sustain the activity, and how hard they are going to pursue it (Dörnyei & Skehan, 2008). Motivation, in addition to aptitude, has served as the most consistent predictor of L2 learning success, producing correlations with language achievement that range between 0.20 and 0.60, with a median value around 0.40. Aside from age of onset (AO) and aptitude, no other potential predictors of L2 learning success have consistently achieved such levels (Dörnyei & Skehan, 2008).

Several models to frame and classify motivation have been used throughout the educational psychology and applied linguistics literature. One of the first such models, proposed by Deci and Ryan (1985), contrasts *intrinsic* and *extrinsic* motivational orientations. According to this model, intrinsic motivation comes from within the individual and as such is directly related to one's identity and sense of well-being. Students are intrinsically motivated when learning is a goal in and of itself and when learners find tasks interesting and challenging. In this sense, the reward is the enjoyment of the activity itself or a feeling of competence (i.e., self-efficacy) in completing the task (Bandura, 1997). Extrinsic motivation, on the other hand, originates from outside the individual. Students are extrinsically motivated when learning is done for the sake of rewards (e.g., grades or praise) that are not inherently associated with the learning itself. Research has shown that intrinsic motivation correlates more closely with language learning success than extrinsic motivation; however, a student's total motivation is most frequently a combination of both extrinsic and intrinsic orientations (e.g., Walqui, 2000).

While Deci and Ryan's (1985) intrinsic-extrinsic dichotomy predominantly focuses on the origins of motivation, Gardner (1985) structured a model of motivation around learners' objectives for engaging in language learning, which fall into two broad categories: *integrative* and *instrumental* orientations. According to Gardner, an integrative motivation orientation reflects a positive disposition toward the target language community and a desire to interact with and become similar to valued members of that community (i.e., to integrate into the community). In contrast, an instrumental orientation posits language learning as primarily associated with the potential and perceived pragmatic gains of target language proficiency, such as getting a better job or a higher salary. Studies have suggested that learners who have a strong desire to integrate into the target language community (i.e., integrative orientation) are both more motivated overall and exhibit higher language attainment than learners who report being more instrumentally oriented, such as for reasons of academic or career advancement (e.g., Gardner & Lambert, 1972).

However, two major limitations arise from the motivational orientation models discussed thus far: the implied framing of motivation as a *static* trait and the limited scope and relevance of integrative motivation in certain learning contexts outside of immigration and study abroad. An integrative orientation has proven far less important in foreign language settings where such social integration in the target language community is virtually impossible (Au, 1988; Crookes & Schmidt, 1991; Oxford & Shearin, 1994; Leaver, 2003). Furthermore, motivation is not, in reality, a static trait but considerably *dynamic*. During the lengthy process of mastering certain subject matters, motivation does not remain constant, but rather is associated with dynamically changing and evolving mental processes, characterized by continuous reappraisal and balancing of the various internal and external influences to which individuals are exposed.

Therefore, Dörnyei (2000) argues that in order to account for the daily ebb and flow of motivation, an adequate model must have a distinct temporal dimension that can accommodate systematic patterns of transformation and evolution through time (Dörnyei, 2000). In an attempt to address the challenge of time in theories of learner motivation, Dörnyei and Ottó (1998) proposed a process-oriented conceptualization of motivation. Their process-oriented approach uses time as an organizing principle thereby offering a natural way of ordering the relevant motivational influences into various distinct stages along a temporal axis. Dörnyei's (2000, 2001) elaborated model of the process-oriented approach to motivation includes three phases: 1. the *pre-actional phase*, associated with *choice motivation* for forming intentions, setting goals, and initiating actions; 2. the *actional phase*, associated with *executive motivation*, for generating and carrying out sub-tasks, conducting ongoing appraisal, and action control; and 3. the *post-actional* phase, which involves the learner's final analysis of the actional process once it has been completed or terminated. Each of these three phases involves different motivational functions and influences.

To address the inherent limitations of Gardner's original conceptualization of integrativeness as a motivating desire to integrate into the target language community, Dörnyei (2005, 2009) developed an updated framework, the L2 Motivational Self System (L2MSS), which draws on the psychological theory of *possible selves* (Markus & Nurius, 1986, 1987). Possible future selves represent individuals' ideas of what they might become, what they would like to become, and what they are afraid of becoming and thereby function as future self-guides that channel and give direction to current motivational behaviors, providing a link between the self-concept and motivation (Markus & Nurius, 1986). Dörnyei and Csizér (2002) reasoned that the process of identification underlying the concept of integrativeness might be better explained as an internal process of identification with a projected future image (i.e., “Ideal L2 Self”) within the person’s self-concept, rather than identification with an external reference group such as the L2 community. Dörnyei’s (2005, 2009) L2MSS consists of three primary constituents in addition to a number of conditions that need to be in place for these constituents to have sufficient motivational potency: 1. the *ideal L2 self*, which concerns a desirable self-image of the kind of L2 user one would ideally like to be in the future such that if a discrepancy is observed between the ideal L2 self and one’s current state, the individual may be motivated to learn a new language or further develop their proficiency in an existing one; 2. the *ought-to L2 self*, which reflects the attributes that one believes one ought to possess in order to meet the expectations of others and to avoid possible negative outcomes; and 3. the *L2 learning experience*, which focuses on the learner’s present experience, covering a range of situated, executive motives related to the immediate learning environment, such as the impact of the teacher, the curriculum, the peer group, and the experience of success.

In sum, the L2 Motivational Self System (L2MSS) suggests three primary sources of motivation to learn another language: the learner’s internal desire to become an effective L2 user, the social pressures to master the L2 coming from the learner’s environment, and the actual experience of being engaged in the L2 learning process (Dörnyei, 2009). L2MSS has become the most prominent language learning motivation theory and therefore the present study utilizes a motivation questionnaire with a Spanish-oriented adaptation of Dörnyei’s L2MSS framework (2005, 2009) to characterize individual learners’ degree and nature of motivation to learn Spanish. Consult the Methodology section for further details on the motivation instrument used in this study.

Bigg’s model (1992), on the other hand, exploits the connection between intrinsic motivation and *deep strategies*, treating motivation in parallel with learning strategies. This model addresses both motivation and learning strategies by categorizing learning strategies into categories relative to a learner’s motivational orientation: 1. *surface strategies* to get a task done with as little personal investment as possible, involving surface processing to complete the task at hand with minimum conceptual effort. As a result, much less information will remain in long-term memory when employing surface strategies because the information has been encountered much less and there is no emotional or cognitive investment (Ehrman, 1996); 2. *achieving strategies* used to succeed in competition and get good marks; and 3. *deep strategies*, which make personal investments in the task through associations and elaboration. The deployment of deep strategies leads to *deep processing*, which is an active process of making associations with material that is already familiar, examining interrelationships within the new material, elaborating the stimulus through associations with it and further development of it, connecting the new material with personal experience, and considering alternative interpretations. Furthermore, during deep processing, the learner may use the new material to actively reconstruct their conceptual frameworks (Ehrman, 1996). In sum, surface processing occurs when there is no emotional investment, whereas deep processing involves a more profound and integrated interpretation and processing of new material. Therefore, processing type must be interrelated to some extent with learning motivations and attitudes; if one feels emotionally engaged with the language and language community, perhaps this fundamentally alters the way the individual processes the target language input thereby leading to deeper, and subsequently more effective, processing. The effect of learner attitudes toward the target language and target language community will be considered next.

Attitudes

Social psychological theories of action argue that attitudes exert a directive influence on people's behavior since an individual's attitude toward a target influences the overall pattern of their responses to the target (Ajzen, 1988; Eagly & Chaiken, 1993). Attitudes related to the target language community exert a strong influence on one's language learning. Evidence for this assertion comes from the observation that few late learners are likely to be successful in learning the language of a low-status community (Dörnyei & Skehan, 2008). Given that native speakers of the target language community are the closest parallels to the idealized L2-speaking self, attitudes towards members of this ethnolinguistic community must be directly related to one's ideal language self-image (Dörnyei, 2009); the more positive our disposition towards the community of speakers, the more attractive our idealized L2 self becomes because people are attracted to others who emulate the person they want to be rather than the person they actually are (Herbst et al., 2003).

Attitudes are largely informed by goals. Gardner (1985) categorized language learners' goals into two broad categories: integrative orientation and instrumental orientation, as previously discussed. An integrative orientation reflects a positive disposition toward the L2 group and the desire to interact with and even become similar to valued members of that community, which is necessarily influenced by the attitudes one holds about the target language community. An integrative motivational orientation is a complex construct made up of three main components: 1. integrativeness, including interest in foreign languages more broadly and attitudes toward the specific L2 community; 2. attitudes toward the learning situation, comprising attitudes toward the teacher and the course; and 3. motivation, made up of motivational intensity, desire to learn the language, and general attitudes towards the act of learning the language. Overall, integrativeness is thought to play a key role in L2 motivation, mediating the effects of all other attitudinal and motivational variables (Gardner, 1985). This idea of *integrativeness* from the self-perspective can be conceived of as the L2-specific facet of one's ideal self; if our ideal self is associated with the mastery of an L2, that is, if the person that we would like to become is proficient in the L2, we can be described in Gardner's (1985) terminology as having an integrative disposition. Therefore, the integrativeness dimension of motivation has been equated with the *Ideal L2 Self*, and this concept was later adopted and adapted by Dörnyei (2005, 2009) in the updated framework, the L2 Motivational Self System (L2MSS), drawing on the psychological theory of *possible selves* (Markus & Nurius, 1986, 1987).

To assess attitudes, the present study utilizes questions taken from the language attitudes section of the Bilingual Language Profile (BLP) questionnaire (Birdsong, Gertken, & Amengual, 2012) regarding one's feelings, identity, and values towards the target language and its associated community of speakers. More information about how language attitudes were measured in the present study via the self-report questionnaire can be found in the Methodology section.

Metacognitive awareness

The concept of metacognition is particularly relevant to adult learners in the instructed context as it refers to the ability to reflect upon, understand, and control one's entire learning process and is comprised of different subcomponents, namely *knowledge about* cognition and *regulation of* cognition (Schraw & Dennison, 1994). The subcomponent of metacognitive knowledge is further comprised of declarative, procedural and conditional knowledge and it is the awareness of these three domains of knowledge that drives overall knowledge about cognition (Schraw, 2009). Because metacognition crucially involves both knowledge about and the regulation of cognition, its development uniquely prepares learners to be more strategic and ultimately perform better than their peers who have less developed metacognitive awareness, as more metacognitively aware individuals are also more able to plan, sequence, and monitor their learning in a way that can directly improve their performance (Schraw & Dennison, 1994).

Adult language acquisition research in the instructed setting has readily shown that metacognitive awareness is positively associated with different measures of performance and learning success. For instance, metacognitive awareness has been shown to compensate for low ability or lack of relevant prior knowledge and to promote the development of autonomous learners (García, 2010). Khodabakhshzadeh, Hosseinnia, and Rahimian (2017) examined the predictive power of three independent learner variables—creativity, metacognition and learning style—on foreign language achievement among Iranian English learners (N=122) and found metacognitive awareness (as assessed by Schraw and Dennison's MAI: Metacognitive Awareness Inventory, 1994) to be the strongest predictor of academic achievement in English, as measured by final course grade. Similarly, Nosratinia, Saveiy, and Zaker (2014) examined self-efficacy, metacognitive awareness, and language learning strategy use as three interrelated learner variables among university students (N=150) learning English as a foreign language and found a significant positive correlation between both self-efficacy and metacognitive awareness (as measured by the MAI, Schraw and Dennison, 1994), and also between metacognitive awareness and language learning strategy use. Other empirical studies examining the effect of metacognitive training interventions on different language learning outcomes have reported a positive effect of metacognitive training on learner autonomy (Kissling & O'Donnell, 2015; Victori & Lockhart, 1995), depth and precision in self-correction (Kissling & O'Donnell, 2015), listening comprehension (Vandergrift & Tafaghodtari, 2010; Kohler, 2002), vocabulary acquisition (Kohler, 2002), and overall self-reported proficiency (Victori & Lockhart, 1995). Furthermore, the relative effect of metacognitive awareness on performance has been shown to be modulated by task demands such that a stronger correlation is observed with self-paced tasks that allow time for planning, and therefore access to explicit knowledge stored in declarative memory, than speeded tasks that place greater demand on implicit knowledge in procedural memory (Black & Tararova, 2020). Therefore, the present study integrates metacognitive awareness as one of the individual variables that may explain some variation in learner performance during varied task demands.

With regards to methodological questions of measurement, there has been ample discussion around how to assess learners' level of metacognitive awareness. Despite this extensive discussion in previous research, few tools exist for assessing metacognitive awareness. In the process of creating a tool for assessing metacognitive awareness specifically as it pertains to adult formal language learning, I reviewed all metacognitive assessment tools freely available on the internet and identified four tools in particular that have been empirically researched and have demonstrated pedagogical applications: The Metacognitive Awareness Inventory (MAI) (Schraw & Dennison, 1994), the Metacognitive Awareness Listening Questionnaire (MSLQ) (Vandergrift, Goh, Mareschal, & Tafaghodtari, 2006), the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich & DeGroot, 1990), and the Strategy Inventory for Language Learning (SILL) (Oxford, 1986). Both the MSLQ and the SILL focus exclusively on learning strategies, as their names indicate, which is one component of overall metacognitive awareness. The MALQ, although providing a more complete assessment of overall metacognitive awareness, focuses exclusively on the listening mode of language learning.

The MAI, on the other hand, provides a more complete assessment of metacognitive awareness in general learning among adults in the instructed context. It includes a total of 52 statements regarding both *knowledge* about cognition and the *regulation* of cognition distributed randomly throughout a written self-assessment questionnaire. Individuals are instructed to mark "true" or "false" for each statement (scoring: "true" response = 1 point; "false" response = 0 points), resulting in a total possible score of 17 points for knowledge about cognition and 35 points for regulation of cognition. Schraw and Dennison empirically tested their MAI among undergraduate students (N=197) in an educational psychology course by comparing individual results from the MAI with test performance scores and confidence judgments about test performance. The results supported the two-factor model of metacognitive awareness (knowledge about and regulation of cognition) and indicated that the MAI provides a reliable initial test of metacognitive awareness in adult students (Schraw & Dennison, 1994). The MAI has subsequently been utilized in numerous other studies that examine the role of metacognitive awareness in both general learning and specifically in language learning within the instructed context (e.g., Khodabakhshzadeh et al., 2017; Nosratinia et al., 2014; García Magdali, 2010; Young & Fry, 2008).

In the present study, the statements to evaluate level of metacognitive awareness were taken from the Metacognitive Awareness Inventory for Language Learning (MAILL) questionnaire developed by the researcher and colleague (Black & Tararova, 2020), which is a shortened adaptation of the Metacognitive Awareness Inventory (MAI) by Schraw & Dennison (1994), applied to the context of learning Spanish in the instructed setting. These statements concern knowledge about cognition and regulation of cognition in the context of learning Spanish in which participants are instructed to self-report how often in a typical week the statements apply to them along a 5-point Likert scale. A full description of how metacognitive awareness was evaluated in this study is provided in the Methodology section.

Metalinguistic awareness

Paradis's neurolinguistic theory of bilingualism (2004, 2009) posits metalinguistic knowledge as one of the four cerebral mechanisms involved in the acquisition and use of language (L1/L2), along with implicit linguistic competence, pragmatics, and motivation. L1 children and early L2 learners engage almost solely in incidental acquisition, which arises through procedural memory and leads to implicit competence or linguistic intuition (Paradis, 2004). Adolescents and adult L2 learners, on the other hand, can no longer build procedural representations to the same extent as children and therefore learn the L2 intentionally while relying on declarative memory, which leads to explicit competence or metalinguistic knowledge and awareness (Paradis, 2004, 2009). To the extent that nonnative speakers have gaps in their implicit linguistic competence, they will compensate by relying more extensively on metalinguistic knowledge and pragmatic aspects of verbal communication in both speaking and understanding (Paradis, 2004). There are considerable individual differences in both nonnative and native language users' linguistic competence (Dąbrowska, 2012; Farmer, Misyak, & Christiansen, 2012; Hulstijn, 2015) that are partly attributable to differences in experience, namely education and print exposure, and partly due to learner-internal factors such as metalinguistic abilities (Paradis, 2009). Brooks & Kempe

(2013) examined adult learner performance with Russian gender and case-marking patterns and found that even under implicit learning conditions, individual differences stemmed from explicit metalinguistic awareness of the underlying grammar. The present study will examine the degree of metalinguistic awareness of Spanish grammatical gender agreement both indirectly through a written grammaticality judgement task and directly through a metalinguistic awareness exit survey at the end of the experimental session (see the Methodology section).

Language use

In addition to motivation, attitudes, metacognitive awareness, and metalinguistic awareness reviewed so far, nonnative adult learners also vary according to the extent and context in which they use the target language. Pavlenko (2002) argues that up until recently, social aspects of nonnative language learning and use have been both under-represented and undertheorized in the SLA literature. Firth and Wagner (2007) go even further as to argue that a distinction between acquisition and use is highly problematic in that both concepts are so tightly interwoven as to be rendered effectively inseparable. Nonetheless, empirical research abounds in recent decades with evidence of the connection between target language use and different measures of language learning. Gardner and Lysynchuk (1990), for example, examine the retention of L2 French skills among high school students after a 9-month absence of instruction, what they referred to as the "incubation period". They found that L2 French proficiency following the incubation period was dependent upon language use during that period and concluded that the extent to which language skills are lost, retained, or improved depends largely upon language use after training has ceased. More recently, De Carli et al. (2015) examined the effect of bilingual language use in Italian-Spanish bilinguals who were classified according to their age of acquisition (AoA) and language use ("intensive" vs. "occasional" language users). Participants performed a pragmatic bilingual test and a battery of cognitive tests. Results indicated that continued language use was a major factor influencing high bilingual proficiency and that this effect of language use was irrespective of AoA. These results largely corroborated the researchers' prediction that intensive language use is a major factor influencing language proficiency for both accuracy and speed of recognition. De Carli and colleagues concluded that the highest levels of proficiency were not associated with AoA, as commonly thought, but rather with intensive language use. As a result, De Carli et al. argue that the Critical Period Hypothesis (CPH) can be questioned in this light as other confounding factors, such as education and language use, could explain the AoA effect commonly detected in age effects research.

Tarone (2007), in a discussion of the sociolinguistic approaches to SLA research, argues that language learners intentionally assert social identities through their L2 in communicating in social context. Therefore, it seems that language attitudes and language use would necessarily be intertwined. This study therefore aims to explore not only the effect of individual variables such as language use, attitudes, and motivation, but also examines any interaction that may occur between these important factors in adult learners. Tarone (2007) points out that although a number of SLA approaches explore different aspects of the relationship between social context, cognition, and L2 use, few studies have examined how language use directly affects the learner's acquisition of specific L2 linguistic forms, as explored in the present dissertation by relating learner performance with grammatical gender in Spanish to their average self-reported weekly use of Spanish both in and outside the classroom.

The multilingual effect

A multilingual turn is under way in SLA research. Researchers are increasingly adopting a focus on diverse linguistic contexts as multilingualism becomes the new norm of applied linguistic and sociolinguistic analysis (May, 2014). Most of the world's population is bilingual or multilingual. However, research on monolingualism has traditionally been prioritized over multilingualism as monolinguals have traditionally served as a sort of benchmark or control by which to compare bi- and multilinguals (Cummins 2007, 2009). Nonetheless, it is essential to recognize that multilingualism is a dynamic cognitive system that is qualitatively different from the cognitive systems of monolinguals (Cummins 2007, 2009). This difference is relevant for language acquisition research since bilingual and multilingual speakers have an empirically demonstrated advantage when it comes to the task of learning a third (L3) or additional (L*n*) language in lexical (e.g., Kaushanskaya & Marian, 2009), phonological (e.g., Tremblay & Sabourin, 2012), phonetic (e.g., Antoniou et al., 2015), and syntactic (e.g., Klein, 1995) domains. Herdina and Jessner (2002) propose that the presence of more than one language system influences the development of all languages represented in the dynamic mental grammar of the individual, including L1, L2, and any additional language. An empirically observed multilingual advantage has been explained in terms of metalinguistic awareness, learning strategies, linguistic repertoire (e.g., Cenoz, 2013), increased language aptitude with experience (Thompson, 2013), affective factors such as increased confidence and reduced anxiety (Dewaele, Petrides & Furnham, 2008), and general changes to the cognitive-linguistic system (Hirosh & Degani, 2018) that are associated with multilingualism during subsequent language learning (L*n*).

Although the terms *bilingualism* and *multilingualism* have been used interchangeably in much research, and have even been equated with one another (see: Bhatia, 2017), for the purpose of the present study, an important distinction is made between bilingual and multilingual language experience; participants with exclusively English-Spanish (nonnative Spanish learners)/Spanish-English (native Spanish speakers) experience are grouped together in the "bilingual" participant group and participants with reported proficiency in additional languages are grouped together in the "multilingual" participant group for the analysis of the possible effect of multilingualism on performance with grammatical gender in Spanish.

Cenoz (2013) argues that the observed multilingual advantage in additional language learning can be largely attributed to multilingual learners' larger linguistic and intercultural repertoires which they can employ to their advantage in L*n* learning. However, this advantage can also be linked to more extensive experience with language learning as a specific set of skills such that multilinguals tend to have more developed and effective language learning strategies, that they have refined over time, than their monolingual peers who are approaching the task of learning an L2. L*n* learners can develop a higher level of metalinguistic awareness on the basis of their previous experience with the task of learning a second language (L2) and their knowledge of two linguistic systems (Kemp, 2007). The idea is that multilingual learners are able to think about language in a more abstract way and regard it as an object (Cenoz, 2013). As a result of this experience, multilinguals have developed a wider range of learning strategies that help them to learn the L3 or L*n*. For example, they seek out more sources of input, make an early effort to use the new language in a meaningful way, and self-direction and an overall more positive attitude toward the language learning task (Cenoz, 2013). Furthermore, language learners with multilingual experience tend to have more grammar- and vocabulary-learning strategies that they use more frequently (Psaltou-Joycey & Kantaridou, 2009; Kemp, 2007) and have also been shown to use qualitatively different learning strategies including memory and compensation strategies (Mitits, 2016) when compared to their bilingual or monolingual peers. Based on the significant increase in both quality and frequency of learning strategies employed by bilinguals and trilinguals of novice L*n* German, Dmitrenko (2017) asserts that there is a "threshold effect" in L3 learning, making it qualitatively different from learning an L2. However, there is some contradictory evidence, such as Mady (2017), who found that immigrant multilingual students in a French immersion program in Canada outperformed their Canadian-born bilingual peers on the French proficiency tests, yet no significant differences were detected between groups for metalinguistic knowledge nor strategy use, suggesting that the multilingual advantage observed is not attributable alone to differences in language learning strategy use or level of metalinguistic awareness.

In addition, the positive influence of the linguistic repertoire mentioned previously has been specifically associated with *language distance* (i.e., typology); in other words, closely related languages would be more useful for multilingual learners attempting to learn an additional language (Ringbom, 2007; Jarvis & Pavlenko, 2008). The degree of typological overlap between the languages of the multilingual speaker may play an important role in the pattern and quantity of transfer when learning a novel language and the effect of language similarity may surface when the structure or concept in question is particularly difficult to learn (Hirosh & Degani, 2018), as is the case of adults attempting to acquire grammatical gender in Spanish. However, there is some empirical evidence that typology between languages is not such a deterministic factor for vocabulary learning (see: Kaushanskaya & Marian, 2009). Therefore, positive transfer effects from other languages known cannot reliably and fully account for a multilingual advantage in novel language learning.

Hirosh and Degani (2018) offer a systematic review of empirical studies (N=33) that examine potential differences between monolingual and multilingual speakers in novel language learning in both children and adults for vocabulary, phonology, grammar, and literacy and categorize the findings of each study according to direct and indirect influences of multilingualism on novel language acquisition. Their thorough review uncovered two main subcomponents of the multilingual advantage, including direct transfer of prior knowledge and prior skills as well as indirect influences that result from a multilingual background, including more general changes to the cognitive-linguistic system. From their findings, Hirosh and Degani (2018) outline their theoretical framework for the influence of multilingualism on novel language learning with two broad divisions: *direct* and *indirect* influence. Knowledge and skills fall under the subcomponent of "direct influence", while linguistic and non-linguistic/executive function fall under the "indirect influence" that constitutes the enhanced cognitive abilities associated with multilingualism. Furthermore, Hirosh and Degani (2018) assert that prior language learning context (i.e., instructed vs. naturalistic) appears to influence the balance between the direct and indirect effects of multilingualism on novel language learning such that formal language learning settings (i.e., classroom) tend to favor more direct multilingual effects (e.g., grammatical knowledge and learning strategies) whereas individuals exposed informally to a multilingual environment tend to rely more on the indirect effects of multilingualism (e.g., inhibition and attention control, ambiguity processing, verbal memory, lexical-semantic network, etc.). According to the framework of Hirosh and Degani (2018) constructed from empirical findings, the effect of multilingualism on L*n* learning touches the knowledge domains of vocabulary, grammar, literacy, accent and phonology, as well as the subdomains of fluency in reading, writing, listening, and speaking. Therefore, the influence of multilingualism appears to be both robust and extensive.

Despite the positive effect of multilingualism reported extensively in empirical studies, Hirosh and Degani (2018) assert that research is still lacking on the differences between monolinguals and multilinguals in learning, in particular, the *grammar* of a novel language. Therefore, the present study is motivated by a notable lack in previous late acquisition research that specifically examines multilingualism as an individual difference that may variably modulate performance with a specific target language (TL) grammatical structure.

2.4 Variationist approach to language acquisition research

Examining variation in SLA

One of the main differences between L1 and L2 acquisition is the significantly increased variability as L2 development is far more exposed to the impact of system complexity than native language development. This variability and complexity in L2 development is reflected in the heterogeneity of the typically limited end state of adult learners' language attainment (Dörnyei, 2009). Nonnative, and particularly late, language acquisition can only be explained, in part, by considering a variety of learner-based and environmental factors, such as the age of onset, motivation level, as well as the amount and nature of input available in different learning contexts. Larsen-Freeman (2012) crucially argues that the difference between language learners is not merely 'noise', but rather a natural part of dynamically emergent behavior of individuals with different orientations, traits, and experiences. In this sense, there is no norm from which individuals deviate but rather variability that stems from the ongoing self-organization of systems of activity (Larsen-Freeman, 2012). From this dynamic and emergent perspective, language learning is not merely the integration of linguistic structures by learners, but also the constant adaptation and enactment of language patterns in the service of meaning-making in response to the properties and corresponding functions that emerge in a dynamic communicative situation (Larsen-Freeman & Cameron, 2008).

Complexity and Dynamic Systems Theory (CT/DST) in SLA

Viewing language as a complex, dynamic system and language use/acquisition as dynamic adaptedness constitutes a way of understanding the change in progress that is a developing system (Larsen-Freeman, 2006: p. 590). Although Complexity Theory (CT) and Dynamic Systems Theory (DST) originated in the physical sciences and mathematics, the theory has come to be embraced by a range of disciplines from biology, organizational development, to epidemiology. CT/DST aims to account for how the interacting parts of a complex system, such as language competence, give rise to the system's (collective) behavior and how such a system simultaneously interacts with its environment. CT/DST inspires us to think differently about languages and language development, especially regarding their complexity and dynamism. In this sense, dynamism is central to language evolution and change, language processing and use, and language acquisition and development. Table 4 summarizes the general principles of complex systems according to Complexity Theory (CT)/Dynamic Systems Theory (DST), as discussed in Larsen-Freeman (2012). In sum, adopting a complexity theory approach shifts our perspective on what exactly merits empirical investigation: it changes what we need to collect as data to understand a complex system such as language, in particular our attitude towards variation, and it changes what we notice in the behavior of systems such that flux and variability signal possible processes of self-organization and emergence. Sudden phase shifts signal important changes and can direct attention to the conditions that lead up to them (Ellis & Larsen-Freeman, 2006; Halliday, 2007; Hult, 2010).

Table 4. Twelve general principles of complex systems according to Complexity Theory (CT) and Dynamic Systems Theory (DST) as presented and discussed in Larsen-Freeman (2012).

CT/DST Principle:	**Description:**
1. open & dynamic	Complex systems are open and dynamic.
2. not in equilibrium	They operate under conditions that are not in equilibrium.
3. contain many interacting elements/agents	Complex systems are systems because they comprise many elements or agents, which interact.
4. change/dynamism is central	The systems adapt both through interaction with the environment and through internal reorganization/self-organization.
5. strength of interactions changes over time	Multiple routes are often possible between components, mediated in different ways.
6. system complexity is emergent.	The complexity of complex systems is not built into any one element or agent, but rather arises from their interaction.
7. nonlinear systems	Because the systems are open, what arises may be in nonlinear relation to its cause; an unexpected occurrence may take place at any time.
8. structure maintained despite changing components	The structure of a complex system is maintained even though its components may change.
9. environment is itself a complex system	The environment in which they operate is part of a complex system.
10. operate over a range of timescales and levels of complexity	Complex systems display behavior over a range of timescales and at different levels of complexity—the latter are nested, one within the other.
11. display chaotic variation	Complex systems sometimes display chaotic variation.
12. iterate	Complex systems iterate: they revisit the same territory again and again, which means that the present level of development is critically dependent on what preceded it.

The iteration of CT/DST most commonly applied to SLA is known as *Emergentism*. The emergentist shift of perspective in SLA provides another way of understanding previously observed characteristics of learner language: language development is not discrete and stage-like but more like the waxing and waning of patterns of development. Certain aspects of the behavior are progressive while others are more regressive. Change can be gradual and it can also be sudden (Larsen-Freeman, 2006). As Selinker argues “a theory of second language learning that does not provide a central place for individual differences among learners cannot be considered acceptable” (Selinker, 1972: p. 213), and intra-learner variation is key in understanding the dynamic system of language development (Larsen-Freeman, 2012).

Examining variation as dependent variable in native and nonnative performance

The concept of *nativeness* and the role of the *idealized native speaker* in questions concerning nonnative ultimate attainment are currently and rightly being questioned in the language acquisition literature. There is growing evidence that native speaker convergence is a myth as there are considerable individual differences in adult L1 speakers' linguistic competence (Dąbrowska, 2012; Farmer et al., 2012; Hulstijn, 2015). Such native speaker differences are attributable in part to differences in language-related experience, in particular level of education and degree and depth in print exposure (i.e., literacy), and also partly attributable to learner internal factors such as statistical learning abilities, intelligence quotient (IQ), and metalinguistic abilities (Andringa & Dąbrowska, 2019). Therefore, these individual variables are relevant for both native speakers and adult learners alike.

The reality of native speaker variation has important implications for SLA research. Although the majority of ultimate attainment studies use highly educated participants as a native control group to which to compare L2 learners (Andringa, 2014), a very different picture emerges from studies that use a native control group that includes lower socioeconomic-status speakers and these studies tend to show more L2 learners performing within the native speaker range (e.g., Andringa, 2014; Dąbrowska & Street, 2006; Hulstijn, 2015). In this sense, there are two principled ways of characterizing native speakers: 1. In terms of *shared basic language cognition* (see: BLC framework, Hulstijn, 2015); and 2. *extended or higher language cognition*. While BLC pertains to the frequent units and constructions that are shared by all native speakers, extended language cognition pertains to the infrequent units and constructions that distinguish native speakers from one another along extralinguistic dimensions such as socioeconomic status, level of education, and print exposure (Hulstijn, 2015). Dąbrowska (2019), for example, investigates to what extent several nonlinguistic measures, such as print exposure and level of education, are predictive of attainment in both native and nonnative speaker competence and to what degree they show overlap. It was found that these predictors of attainment are largely similar for native and nonnative speakers, although not necessarily equally weighted for both speaker groups. Therefore, variation should be examined as a dependent variable in both native and nonnative speaker groups in order to elucidate and more accurately represent the dynamics of language use between and among individuals as language grows and organizes itself in a dynamic and organic way as a complex, emergent system (Larsen-Freeman, 2006).

Methodological principles in researching SLA from a CT/DST perspective

Adopting a complexity theory perspective in SLA necessarily shifts our view of what seems to need empirical investigation, in particular, the role of context and environment (Larsen-Freeman & Cameron, 2008). A CT/DST perspective shifts what we need to collect as data to understand a complex system and critically concerns our attitude towards *variation*, thereby changing what we know in the behavior of systems (Larsen-Freeman & Cameron, 2008). Larsen-Freeman and Cameron (2008) outline methodological principles to orient language development research from a CT/DST perspective. Such principles include: to be ecologically valid by including context (e.g., context of acquisition) as part of the system under investigation; to honor complexity by avoiding reductionism and including as many factors as possible that might influence the system; treat associations and relationships between variables as nonlinear, multivariate, and interactive; rethink units of analysis, identifying collective variables that characterize the interaction among multiple elements in a system; and consider variability as central by investigating both stability and variability in order to understand the developing system. For the purpose of the present study, the following CT/DST guidelines were adopted in the manner outlined below:

> ***1. Context of acquisition is considered*** as an independent, potentially explanatory, variable by grouping participants for analysis as native and nonnative speakers;
>
> ***2. A wide variety of factors are considered***, including individual-internal variables (motivation and attitudes, metacognitive awareness, metalinguistic awareness, proficiency, and degree and type of multilingualism), linguistic variables (noun gender class, morphological marking, and domain of agreement), and task variables (stimuli modality and knowledge type favored) that may variably draw upon the developing linguistic system of the individual thereby affecting linguistic performance;
>
> ***3. Potential interactions*** between independent variables are examined through correlation matrices;
>
> ***4. A multivariate approach*** to modeling predictors of language performance is adopted for inferential data analysis;

5. Inclusive linguistic performance is evaluated both by measures of central tendency in accuracy (i.e., mean) and by measures of dispersion (e.g., standard deviation, variance, standard error) in order to treat intra-speaker variation as a dependent variable on the same level as accuracy. Furthermore, treating variation as a dependent response variable allows us to draw some principled conclusions regarding the knowledge source and processing type evident in speaker and learner performance since explicit knowledge is posited to be much more variable than implicit knowledge (R. Ellis, 2005).

2.5 From Emergentism to generativist accounts of language acquisition

Up until now, we have examined the language acquisition literature from emergentist, cognitivist, and variationist perspectives. However, most of the work specifically on the acquisition of grammatical gender has been done within the generativist approach to language acquisition research, based on the principles of generative and Universal Grammar (UG), the innate biological component of human language faculty (see Chomsky, 1986). The main interest of researchers within the generativist framework is to understand how language is represented in the mind and how those representations are formed through the acquisition process. Therefore, implicit linguistic knowledge is central to a generativist approach. Generativists argue that there is an innate and biological component to human language called Universal Grammar (UG) for native language acquisition; however, the late acquisition research debate centers around whether and to what extent late learners still maintain access to UG for subsequent language acquisition, particularly post-puberty (i.e., post neurological maturity). Research from a generative standpoint on nonnative/late language acquisition has been centered around two theoretical debates regarding the abstract representation of the mental grammar as it relates to access to UG, that is, whether late acquisition is still UG-constrained. The *representational deficit* account argues that adult language learners are forced to rely on *explicit* or *general* learning mechanisms to acquire another language post-puberty since they do not maintain access to the implicit and linguistically specific learning mechanisms of UG. The *full access* account, in contrast, argues that adults maintain (some) access to the implicit, innate, and language-specific learning mechanisms of UG. Therefore, although the present dissertation largely adopts an emergentist and cognitivist approach in its orientation and eventual interpretation and explanation of the findings, most of the research on the late acquisition of grammatical gender on which this study is based has been carried out by researchers oriented in the generativist framework. We will revisit later on the deficit and full-access generativist accounts as they pertain specifically to the debate on whether the acquisition of grammatical gender is governed by UG in late language learning. First, we will examine gender as a typological feature of language and then we will consider previous empirical research and theoretical explanations grounded in the generativist framework for findings of late adult learner performance with grammatical gender in Spanish.

2.6 Grammatical gender and its acquisition

In the present study, grammatical gender was chosen as the target structure for analysis of linguistic performance for two principled reasons: 1. Grammatical gender is particularly problematic for adult learners of Spanish, leading to persistent and fossilized errors despite overall increasing proficiency (e.g., Fernández-García 1999, McCarthy 2008, Montrul et al., 2008); and 2. Grammatical gender presents several inherent features—namely morphological marking, noun gender class, and domain of agreement—with which to examine differentiated performance not only overall, but also with specific components of the morphosyntactic structure in question as this may be relevant when considering how both native and nonnative language users may differentially process linguistic input.

Grammatical gender: Spanish

When we examine grammatical gender as a typological feature of language, we find that it is present in some form in approximately twenty-five percent of world languages (Corbett, 1991). Grammatical gender is a specific form of a larger noun class system that allows for the formation of agreement between nouns and other linguistic structures such as adjectives, articles, pronouns and/or verbs (Corbett, 1991). World languages that feature grammatical gender include Spanish, French, Italian, Portuguese, Russian, German, Arabic, and Hindi, among others. Common systems of gender division include masculine/feminine, masculine/feminine/neuter, animate/inanimate, and common/neuter.

Grammatical gender in Spanish, like most Romance languages, is a binary system in which all nouns are assigned as masculine or feminine. Although gender assignation is a lexical property of nouns, grammatical gender is realized at the syntactic level in which there must be agreement between a noun and its determiner and modifier(s), thus resulting in two domains of grammatical gender in Spanish: assignment (lexical) and agreement (syntactic) (Alarcón 2009; 2011). Contrary to animate nouns in which gender is semantically motivated (e.g. *el doctor* "the-masc. doctor-masc.", *la doctora* "the-fem. doctor-fem."), inanimate nouns in Spanish have a gender assignment that is *not* semantically motivated and therefore is purely grammatical in nature (e.g., *el libro* "the-masc. book", *la silla* "the-fem. chair"), making their classification arbitrary (Montrul et al., 2008). Finally, noun morphology is also a relevant linguistic feature of grammatical gender in Spanish in which nouns can also be classified based on their level of morphological *transparency* for gender. Most nouns in Spanish follow a canonical or prototypical pattern in which their respective morphology reveals their grammatical gender class, such that masculine nouns tend to end in /-o/ and feminine nouns tend to end in /-a/ (e.g., *el libr**o*** "the-masc. book"; *la sill**a*** "the-fem. chair") (Montrul et al., 2008). This morphologically prototypical group of nouns can be described as canonical or overt, with regards to their morphology. There are other nouns that do not follow this prototypical pattern and therefore can be classified as non-canonical or non-overt because their morphology does not directly reveal information about their grammatical gender class. These morphological variants include nouns that end in /-e/ (e.g., *el puente* "the-masc. bridge") and in consonant (e.g., *la flor* "the-fem. flower"). There is also a subclass of non-overtly marked nouns which can be classified as exceptional in that, within this category, masculine nouns end in /-a/ (e.g., *el problem**a*** "the-masc. problem") and feminine nouns end in /-o/ (e.g., *la man**o*** "the-fem. hand"), thus contradicting the prototypical pattern (Montrul et al., 2008; Alarcón, 2011; Foote, 2015). Nonetheless, according to Teschner and Russell (1984), 99.87% of all nouns that end in /-o/ are masculine and 96.30% of all nouns ending in /-a/ are feminine in the official *Diccionario de la Lengua Española*. Therefore, these exceptional cases are indeed rare and have therefore been excluded from the present study.

Native and nonnative acquisition and processing of grammatical gender in Spanish

Acquiring the parameter of grammatical gender in language learning (L1 and L2) involves both the lexical level—by learning the meaning of a noun together with its inherent gender feature (gender assignment)—and the syntactic level—by learning to establish agreement between the noun and other elements in verb and noun phrases (gender agreement). Thus, it can be argued that learners of Spanish need to acquire the nominal feature of gender in their implicit mental grammar system *before* making valid form-function mappings (Alarcón, 2011). Grammatical gender is acquired at a relatively early age; children as young as three-years-old have been shown to actively make use of grammatical gender cues (Lew Williams & Fernald, 2007; Pérez Pereira, 1991; Hernández Pina, 1984). Nonetheless, empirical research also shows that in L2 acquisition of Spanish, grammatical gender is persistently problematic and therefore is acquired relatively late in the L2 learning process (McCarthy, 2008; Fernández-García, 1999), and is also vulnerable to fossilization, or incomplete/different acquisition, resulting in permanent non-target-like forms, despite increasing overall L2 proficiency. Even advanced L2 learners display persistent errors with gender agreement, primarily in their spontaneous oral production (Montrul et al., 2008).

English does not have purely grammatical gender, only semantic (biological) gender, and thus English nouns do not display morphological marking for gender. Therefore, adult English-speaking learners of L2 or L*n* Spanish must acquire a new parameter setting for nouns in their developing linguistic system and then subsequently learn to map these newly acquired gendered nouns onto syntactic structures to produce agreement in noun phrases (Montrul et al., 2008). The essential question then becomes whether adult learners of Spanish can acquire a parametrized functional feature (i.e., grammatical gender) that is not instantiated in their L1 (e.g., English) (Spino-Seijas, 2017). Empirical evidence suggests that adult learners whose L1 lacks grammatical gender generally find gender agreement more difficult to master than learners whose L1 has grammatical gender (Montrul et al., 2008), although there is some empirical evidence to the contrary (see: Bruhn de Garavito & White, 2002). Given an L1 in which grammatical gender is not instantiated, such as English, L2 Spanish interpretation and production errors can be linked to the linguistic variables involved in grammatical gender assignment and agreement: gender class (masculine or feminine), domain of agreement (article-noun, noun-adjective), and morphology (overt/canonical or non-overt/noncanonical) (Foote, 2015; Gamboa, 2012; Alarcón, 2011; Montrul et al., 2008). The results of most previous research examining adult L1 English-speakers acquiring L2 Spanish suggest that production errors are more common for feminine nouns than for masculine, errors in agreement (syntactic level) are more frequent than errors in assignment (lexical level), and that errors with overtly marked nouns are much less common than errors with nouns that display non-overt or exceptional morphology (see: Spino-Seijas, 2017;

López Prego, 2015; Gamboa, 2012; Alarcón, 2011; Montrul et al., 2008; Bruhn de Garavito & White, 2002; Bruhn de Garavito, 2007).

Adult learners of Spanish exhibit persistent errors particularly in their oral production of Spanish grammatical gender that can be immediately related to the linguistic variables previously mentioned. Nonetheless, the question remains of *why* these linguistic variables present persistent difficulties for late learners acquiring Spanish gender. Some researchers suggest that there is a *representational deficit* or a *maturational constraint* (also called the "learnability problem") for adult language acquisition (e.g., the Fundamental Difference Hypothesis, Bley-Vroman, 1989, 1990), while others support the view of *full-access* in which errors are *not* the result of a fundamental representational deficit in the mental grammar of the adult learner, but rather the result of a "mapping problem" in nonnative language production, that is, a computational difficulty during online processing of L*n* (Spino-Seijas, 2017; López Prego, 2015; Alarcón, 2011; Montrul et al., 2008). Full access accounts of adult (nonnative) language acquisition often cite the Missing Surface Inflectional Hypothesis (Prévost & White, 2000), which claims that adult language learners do, in fact, have the feature *gender* represented in their respective mental grammars at an abstract syntactic level, but that gender errors still occur due to an assembly or production problem (Gamboa, 2012; Montrul et al., 2008). In contrast, representational deficit accounts of adult language acquisition often relate their results to a hypothesized *critical period*, which has been demonstrated empirically (by several historic cases of feral children, such as the case of "Genie") for L1 acquisition (Penfield & Roberts, 1959; Lenneberg, 1967), yet its implications for adult nonnative acquisition remain less straightforward and may only be relevant for the acquisition of *implicit* linguistic competence (DeKeyser, 2000; Paradis, 2004). According to Bley-Vroman's Fundamental Difference Hypothesis (FDH: 1989; 1990), adult language learners are forced to rely on *explicit* or *general* learning mechanisms to acquire another language post-puberty and do not maintain access to the implicit and linguistically specific learning mechanisms of Universal Grammar (UG). In accordance with Bley-Vroman's FDH, McCarthy (2008) proposes the Morphological Underspecification Hypothesis (MUSH), which also supports the representational deficit view of nonnative language acquisition, arguing that grammatical gender errors may be more common with one gender (such as feminine in Spanish) than the other (e.g., masculine) due to the overgeneralization of a default form in the mental grammars of those whose L1—or prior linguistic repertoire more broadly—does not have grammatical gender already instantiated. The implication of the MUSH for nonnative Spanish acquisition by adult English-speakers, and other speakers of non-gendered languages, is that the masculine form is often treated as a default and is therefore overextended to feminine nouns, resulting in higher gender assignment error rates on feminine nouns, which has been instantiated by L2 Spanish acquisition studies (Montrul et al., 2008; Gamboa, 2012). However, we must be cautious in our interpretation of "default" as an explanation for learner errors as native speakers too have been found to demonstrate a default form for gender assignment on unknown or nonce words (Beatty-Martínez & Dussias, 2019; Pérez-Pereira, 2011; Eddington & Hualde, 2008; Eddington, 2002).

Within the deficit vs. full-access generativist debates of late language acquisition, it is important to consider the role of target language proficiency as most of these claims pertain to ultimate attainment or end-state. White, Valenzuela, Kozlowska-Macgregor, & Leung (2004) investigated how adult L1 French and L1 English speakers acquire gender and number agreement in Spanish. Their results showed significant effects of proficiency such that beginner learners differed significantly from the native speaker control group, but crucially, the advanced and intermediate learner groups did *not* significantly differ from their native speaker peers on an oral production task and an interpretation task using pictures contrasting number and gender agreement. Furthermore, no significant effects of L1 (English vs. French) were detected nor for prior exposure to another gendered language. Therefore, these findings provide evidence that at more advanced levels of Spanish proficiency, it is possible for late learners to mirror native speaker linguistic performance, providing evidence that late learning may also be UG-constrained. Nonetheless, these findings can only be related to observable linguistic behavior and not the underlying mechanisms of acquisition. Perhaps the nature of the task and target stimuli would also play a role in distinguishing between learners of different proficiency levels and native speakers, as we examine through our investigation of task effects per proficiency level in the present study.

In terms of the theoretical distinction between the lexical and syntactic domains of grammatical gender, Grüter et al. (2012) found that assignment errors (as assessed by determiner choice) were ten times more frequent than agreement errors among L2 learners, which also supports the results of Alarcón (2011) who reported a 63% error rate among HL speakers and an 87% error rate among L2 learners on the assignment of non-overtly marked feminine nouns during an oral task compared to agreement error rates below 40% for both groups; however, these results contradict other empirical studies in which both L2 learners and HL speakers performed better with the lexical domain of grammatical gender (i.e., assignment) than they did with the syntactic domain (i.e., agreement) (Spino-Seijas, 2017; López Prego, 2015; Gamboa, 2012; Montrul et al., 2008). It should be noted, however, that there has been some debate as to how to assess the assignment or lexical domain of grammatical gender as a separate (yet intricately related) domain from the agreement or syntactic domain (see: Grüter et al., 2012; Alarcón, 2011; Montrul et al., 2008). Although grammatical gender is conceptualized in theory as encompassing both lexical and syntactic domains, in research, this distinction has been deemed problematic to tease apart. As Montrul et al. (2008) explain, the determiner used is often thought to reflect lexical assignment since research shows that monolingual children appear to use the gender marking on the determiner to predict the gender of the following noun (Carroll, 1989). Grüter et al. (2012), for example, found that L1 speakers were efficient at using the gender on the determiner as a cue for what was coming, and this was also true for L2 learners in the case of unfamiliar words. Nonetheless, Montrul et al. (2008), Alarcón (2011), and others concede that the distinction between agreement and assignment in grammatical gender represents a methodological obstacle best mitigated by discussing these two domains in theory while operationalizing the distinction of "domain of agreement" (rather than absolute domain) by comparing accuracy scores on determiners and adjectives, rather than making the theoretical jump to assume that determiner use solely represents assignment and adjective use solely represents agreement. Error analysis provides important insight into the domain source of errors made by adult learners. Montrul et al. (2008) and Alarcón (2011) categorize errors as pertaining to gender assignment if both the article and the adjective are incorrect, and conversely, as a gender agreement error when only one of the two is incorrect. The problem arises in the attribution of errors such as **la puente rojo*, "the(FEM) bridge(MASC) red(MASC)" (the red bridge), since there appears to be agreement between the noun and the adjective, but a lack of agreement between the determiner and the noun. Such an error could be attributable to both the syntactic domain (i.e., lack of agreement between the determiner and the noun) as well as the lexical domain (i.e., incorrect gender assignment on the noun). Nonetheless, as previous research has shown (e.g., Grüter et al., 2012;

Alarcón, 2011; Montrul et al., 2008), there exists a need to distinguish between these two domains of agreement (determiner, adjective) since nonnative adult learners have shown significant differences in their accuracy scores with these two components of grammatical gender in Spanish. Therefore, for the purpose of this study, accuracy scores on determiners and adjectives are analyzed and compared as they both comprise the two "domains of agreement" of grammatical gender.

In addition to considering the linguistic features of grammatical gender and the two main theoretical viewpoints—namely, *representational deficit* and *full access*—that seek to explain why these linguistic variables result in learner errors, we will now examine past research on different *learner* variables that may influence knowledge and production of grammatical gender in Spanish, namely, age of acquisition, nature of input, context of learning, proficiency level, and working memory. Throughout the years of SLA research, many studies have focused on comparing adult L2 learners to native (monolingual) speakers, as a baseline (see: Spino-Seijas, 2017; Foote, 2015; López Prego, 2015; Gamboa, 2012; Grüter et al., 2012; Sagarra & Herschensohn, 2010, 2013; Alarcón, 2009), demonstrating that native speakers perform at ceiling in both grammatical gender assignment and agreement, independently of task modality, whereas late learners exhibit greater variability in their interpretation and production of gender features and are more affected by task modality, consistently performing better in written interpretation than in oral production, yet also consistently and significantly lower than native speaker controls. However, other recent research on grammatical gender acquisition has examined the role of age of acquisition, quality and quantity of input, and learning context by comparing heritage learners (HL) of Spanish to adult English-speaking L2 learners in the US university context, as their ages of initial exposure, type of input, and learning context differ, despite both groups' dominant language being English (Alarcón, 2011; Montrul et al., 2008). Task modality, that is, written interpretation/recognition vs. oral production, has served as a sort of lens through which learner variables, particularly those that distinguish HL speakers from L2 learners, can be more closely examined. As previously mentioned, native speaker controls tend to perform at ceiling regardless of task modality, yet an effect of modality is observed in non-dominant speakers (HL) and learners (L2/L*n*) of Spanish such that learners tend to perform better than their HL peers on written interpretive tasks, which involve *offline* processing that favors *explicit* grammar knowledge; HL speakers, in contrast, tend to perform better than their L2 learner peers on oral production tasks, which involve *online* processing of language that favors *implicit* grammatical knowledge. Montrul et al. (2008) and Alarcón (2011) provide evidence that the HL and L2 learner participants in their respective studies were both dominant in English and had been matched for Spanish proficiency. Therefore, modality (task) effects must reflect differences in age of exposure to Spanish (HL exposed since birth; L2 learners exposed post-puberty), the nature of Spanish input (naturalistic and oral for HL and more formal and written for L2), and the learning context (home/informal in HL vs. classroom/formal in L2). Both Alarcón (2011) and Montrul et al. (2008) conclude that Spanish gender agreement is acquirable *irrespective of the age* of acquisition and the status of the gender feature in the learners' L1 (English) due to the fact that HL speakers were also found to have problems with gender agreement and assignment, despite very early exposure to the target language. Therefore, both HL speakers and nonnative (L2) learners know something about grammatical gender in Spanish, but, as Montrul et al. (2008) conclude, such knowledge might be stored, represented and reproduced differently in the mental grammars of HL speakers and L2 learners and, furthermore, incomplete (Montrul, 2016) or different (Kupisch & Rothman, 2018) acquisition observed in these two populations may be due to distinct causes: variable and insufficient input for HL speakers and access to different language learning mechanisms for adult L2 learners. Alarcón (2011) argues that post-critical period learners (i.e., adult) are more susceptible to computational deficits observed in spontaneous oral production in gender agreement than those who acquire the target language since birth, such as in the case of HL speakers.

Several contradictions and gaps still remain from previous research on the acquisition of grammatical gender. For example, in her study comparing the response times of native Spanish speakers (n=22) and L2 Spanish learners (n=139) on a sentence-completion task, Alarcón (2009) found that L2 learners do *not* appear to process grammatical gender morphologically, whereas other studies (Foote, 2015; Montrul et al., 2008) have shown that L2 learners actually perform better in both comprehension and production tasks with nouns that have overt morphological marking for gender (i.e., canonical nouns), thus indicating that L2 learners do in fact use noun morphology as a cue when processing and learning grammatical gender. With regards to gender class (masculine, feminine), López Prego (2015) found that L2 learners of Spanish were more sensitive to feminine nouns, while other studies have found that L2 learners tend to overgeneralize the masculine and erroneously apply it to feminine inanimate nouns (Gamboa, 2012; Montrul et al., 2008). Based on contradictions in previous research, this study closely examines the role that noun morphology (overt, non-overt), domain of agreement (determiner, adjective), and noun gender class (masculine, feminine) may play in both native and nonnative performance with Spanish grammatical gender.

Variability in L1 and L2 processing: the myth of the uniform L1 speaker

The assumption of a uniform and highly efficient native speaker is a view that has long been challenged in the L1 processing literature due to compelling evidence of individual variation in native speakers (e.g., Beatty-Martínez & Dussias, 2019; Caplan & Waters, 1999; Farmer et al., 2012; Just & Carpenter, 1992; Pakulak & Neville, 2010; Tanner & Van Hell, 2014). These studies provide evidence that even when processing features of language that belong to the same "natural class", native speakers can exhibit sensitivity attuned to distributional patterns of language use (Beatty-Martínez & Dussias, 2019). Variability has also been observed in native speaker processing of grammatical gender and differential behavior of masculine and feminine gender has been observed in that the masculine is commonly treated as a default form in both native and nonnative performance (e.g., Beatty-Martínez & Dussias, 2019). For example, Spanish-speaking children are more likely to treat the masculine noun class as default by assigning masculine gender to nouns with irregular phonological cues (Pérez-Pereira, 2011). Adult native speakers also have been shown to favor masculine over feminine when assigning gender to determiners of unknown nouns, loanwords, and hermaphroditic feminine nouns (Eddington, 2002; Eddington & Hualde, 2008). In this sense, several parallels can actually be established between native and nonnative Spanish speakers concerning variability in grammatical gender assignment. Therefore, using morphological cues for gender assignment and treating the masculine as the default form are not phenomena unique to late bilinguals and L2 learners but rather appear to constitute general linguistic processing strategies. The inherent variability in native speaker processing is, therefore, an important factor when explaining purported deviations from the "native norm" reported in nonnative and learner populations. It is for this principled reason that the present study aims to compare both native and nonnative speaker performance in terms of accuracy and variation under different task demands and across the various linguistic variables inherent to grammatical gender processing in Spanish.

2.7 Summary of previous research and aims of this study

In summary, *implicit* knowledge involves intuitive awareness of linguistic norms, procedural-type knowledge of rules and fragments, is systematic, accessed by means of automatic processing during fluent language use, is nonverbalizable, and is most readily and consistently learnable within a critical or sensitive period. *Explicit* knowledge, in contrast, involves conscious awareness of linguistic norms and therefore is less systematic, features declarative-type knowledge of grammatical rules and fragments, is anomalous and variable, is accessed by means of controlled processing requiring some degree of planning, is verbalizable, and is learnable at any age. Both generativist and cognitivist accounts of language acquisition acknowledge that linguistic competence crucially comprises implicit knowledge such that language acquisition is evident in what language users know *intuitively*. Furthermore, the overwhelming consensus in late acquisition research is that nonnative adult acquisition by implicit means alone is limited in its success as it is believed that additional attention is necessary for the relevant associations to be learned and this can be achieved through form-focused instruction that recruits learners' explicit conscious processing. Therefore, explicit learning, and at least some explicit instruction, seems to be necessary to reach target-like norms in late nonnative language development (but see Schwartz, 1993 for a view to the contrary).

Language acquisition, representation, and processing are all tuned to varying degrees by frequencies in the input and previous research has uncovered significant positive correlations between accuracy and frequency, which can be measured as both *type* and *token* frequency. Furthermore, frequency-regularity interactions have been detected such that when a structure is less systematic (i.e., contradicts the canonical pattern), frequency effects become more pronounced, although the effect size of frequency decreases as language learning progresses and formal instruction can modulate natural frequency effects in grammar learning.

Different memory systems are necessarily involved in the implicit-explicit dichotomy of language knowledge and processing. *Declarative* memory supports the learning of general facts and knowledge whereas *procedural* memory supports motor and sequential skills. Furthermore, procedural memory consists of *implicit* knowledge since the knowledge contained therein is difficult to verbalize and access via introspection. In terms of language functions, declarative memory is involved in the acquisition of the mental lexicon whereas procedural memory is involved in the acquisition of the mental grammar. Ullman's declarative-procedural (DP) model formalizes this distinction for language development. Empirical research has subsequently validated the DP Model in L2 learning as behavioral measures of *declarative* memory have been found to predict grammar learning only in early language training whereas measures of *procedural* memory predicted grammar abilities only in later phases of language training (i.e., at higher proficiency levels).

With regards to language processing type and the operationalization of the implicit-explicit distinction, cognitively *online* tasks are those which are more spontaneous, more complex, and that do not allow for planning time and thereby require the use of implicit language knowledge stored in procedural memory while *offline* tasks allow the language learner time to tap into their explicit knowledge stored in declarative memory. Empirical research has corroborated this distinction as higher accuracy rates have been observed during offline tasks than during online tasks in nonnative adult learners. In order to operationalize this distinction, R. Ellis (2005) proposes a set of criteria to be associated with each task type: in order to solicit *implicit* language knowledge, a task should require a response according to "feel" or intuition, involve a time pressure (i.e., speeded), be focused on meaning (*not* structure), and *not* require metalinguistic knowledge in order to successfully perform. In contrast, in order to encourage or condition the use of *explicit* language knowledge, a task should solicit a response according to rules, have no time pressure (i.e., self-paced), be primarily focused on form, and encourage the use of metalinguistic knowledge.

With regards to task demands, we know that different performance tasks are likely to induce late learners to draw differentially on their implicit and explicit knowledge. Task conditions can be manipulated via a time constraint as well as through the modality in which the task stimuli are presented. Psychological research on stimuli modality differences suggests that verbal material presented aurally and visually is processed in different parts of the memory system and by different mechanisms; the written modality enables more elaborate and more accurate language production than the aural modality and is also perceived as less difficult by literate adult language learners. Furthermore, written input may free up attentional resources to notice, attend to, and potentially become more aware of target features in input. These claims have been corroborated by empirical research demonstrating higher accuracy rates and faster reaction times with written as opposed to aural grammaticality judgement tasks. Imposing a time constraint is another way in which the task conditions can be manipulated in order to condition the use of different knowledge and memory systems. When learners are able to self-pace their performance during a task, they can plan and thereby compensate for processing limitations and increased cognitive demands, enhancing both the accuracy and complexity of their language output. However, under an imposed time constraint, the language user is largely limited to their implicit linguistic competence stored in procedural memory during a cognitively online task.

With regards to individual differences that impact performance and attainment, we now know that adult SLA, or late nonnative language acquisition more broadly, is characterized by large and varied individual differences. That is, past the hypothesized critical period for language learning, individual learner variables acquire greater importance in predicting performance and ultimate attainment. In the present study, the individual learner differences examined include motivation, attitudes, metacognitive awareness, metalinguistic awareness, language use, as well as prior linguistic repertoire and multilingualism.

Motivation concerns both the direction and magnitude of human behavior and therefore is responsible for why a person decides to do something, how long they are willing to sustain the activity, and how hard they are going to work to pursue it. Dörnyei (2005; 2009) proposed the L2 Motivational Self System (L2MSS) that draws on the psychological theory of possible selves and suggests three primary sources of motivation to learn another language: 1. the learner's internal desire to become an effective L2 user; 2. the social pressures to master the L2 coming from the learner's environment; and 3. the actual experience of being engaged in the L2 learning process.

Social psychological theories of action argue that attitudes exert a directive influence on people's behavior and attitudes related to the target language community exert a strong influence on one's language learning process. An integrative orientation reflects a positive disposition toward the target language group and the desire to interact with and even become more similar to valued members of that community, which is necessarily influenced by the attitudes one holds about the target language community, toward the learning situation itself (attitudes toward the teacher and the course), and general attitudes one holds toward the act of learning the language.

Regarding the role of learner awareness in the nonnative language acquisition process, this study examines both metacognitive awareness of the learning process and metalinguistic awareness of the target structure. The concept of metacognition is particularly relevant to adult learners in the instructed context as it refers to the ability to reflect upon, understand, and control one's learning process and is comprised of both knowledge about and regulation of cognition. As such, metacognitive development uniquely prepares learners to be more strategic and ultimately perform better as more metacognitively aware individuals are also more able to plan, sequence, and monitor their learning in a way that directly improves their performance in the target language.

With regards to the role of metalinguistic awareness as an individual learner factor, we know that adult nonnative language learners can no longer build procedural linguistic representations to the same extent as children and therefore are obligated to learn language in a more intentional manner relying more heavily (particularly at the onset of language learning) on declarative memory, leading to explicit competence in the form of metalinguistic knowledge about the target language. In other words, to the extent that nonnative speakers have gaps in their implicit linguistic competence, they will compensate by relying more extensively on their developing explicit awareness of language structure, that is, metalinguistic awareness.

In addition, recent empirical research abounds with evidence of the connection between target language use and different measures of language learning. The extent to which language skills are lost, retained, or improved depends largely upon language use both during and after instruction and continued language use has been identified as a major predictive factor of target language proficiency in terms of both accuracy and speed of recognition, largely irrespective of age of acquisition.

Beyond our consideration of typical individual learner factors such as motivation, attitudes, metacognitive and metalinguistic awareness, and language use, this study also crucially examines the effect of multilingualism on novel language learning. A multilingual turn is under way in SLA research as researchers are increasingly adopting a focus on diverse linguistic contexts with multilingualism becoming the new norm. Multilingualism is now considered a dynamic cognitive system that is qualitatively different from the cognitive systems of monolinguals. Furthermore, multilingual learners have an empirically demonstrated advantage over their monolingual and bilingual counterparts when it comes to learning a third (L3) or additional (L*n*) language in lexical, phonological, phonetic, and syntactic domains. Nonetheless, research on the differences between monolingual, bilingual, and multilingual language learning is still lacking, especially with regards to the learning of a particular grammar structure.

Along with considering individual differences, including the incidence of multilingualism in adult language learners, we acknowledge a variationist approach to language acquisition research, which asserts that one of the main differences between L1 and L2 acquisition is the significantly increased variability. Therefore, differences between late language learners should not merely be seen as ‘noise’, but rather a natural part of dynamically emergent behavior of individuals with different orientations, traits, and experiences. Furthermore, viewing language as a complex, dynamic system and language use and acquisition as dynamic adaptedness constitutes a way of understanding the change in progress that is a developing system. Adopting a complexity theory approach shifts our perspective and attitudes toward variation such that variability signals important changes and can direct attention to the conditions that lead up to them. More specifically, intra-learner variation is key in understanding the dynamic system of language development. There is growing evidence that native speaker convergence is a myth as there are considerable individual differences not only in and among nonnative speakers, but also among adult native speakers’ linguistic competence. Therefore, variation should be examined as a dependent (response) variable in both native and nonnative speaker groups in order to elucidate and more accurately represent the dynamics of language use between and within individuals. In sum, a Complexity Theory/Dynamic Systems approach changes what we need to collect as data and critically concerns our attitude towards *variation*; it changes what we know in the behavior of systems such that flux and variability signal possible processes of self-organization and emergence. Treating variation as a dependent response variable allows us to draw some principled conclusions regarding the knowledge source and processing type evident in speaker and learner performance since explicit knowledge is posited to be much more variable than implicit knowledge.

Finally, with regards to grammatical gender and its acquisition, research has shown that grammatical gender is particularly problematic for adult learners of Spanish, leading to persistent and fossilized errors despite overall increasing proficiency. Grammatical gender as a morphosyntactic structure presents several inherent features by which to examine differentiated performance, including the effect of morphological marking (overt vs. non-overt), noun gender class (masculine vs. feminine), and domain of agreement (Det.-N vs. N-Adj.). Grammatical gender is a rather common typological feature and is present in some form in approximately 25% of world languages. It is considered a specific form of a larger noun class system that allows for the formation of agreement between nouns and other linguistic structures. Grammatical gender in Spanish, like most Romance languages, is a binary system in which all nouns are assigned as masculine or feminine. Although gender assignation is a lexical property of nouns, grammatical gender is operationalized at the syntactic level in which there must be agreement between a noun and its determiner and modifier(s), thus resulting in two domains of grammatical gender in Spanish: assignment (lexical) and agreement (syntactic).

Acquiring the parameter of grammatical gender in language learning (L1 and L*n*) involves both the lexical level—by learning the meaning of a noun together with its inherent gender feature (*gender assignment*)—and the syntactic level—by learning to establish agreement between the noun and other elements in verb and noun phrases (*gender agreement*). The results of most previous research examining adult L1 English speakers acquiring L2 Spanish suggest that errors are more common for feminine nouns than for masculine, errors in agreement (syntactic level) are more frequent than errors in assignment (lexical level), and that errors with overtly marked nouns, that is, nouns whose morphology follows the prototypical pattern for gender marking in

Spanish, are much less common than errors with nouns that display non-overt or exceptional morphology. To explain these tendencies, some researchers suggest that there is a representational deficit or a maturational constraint (the "learnability problem") for adult language acquisition, while others support the view of full access to UG in which errors are not the result of a fundamental representational deficit in the mental grammar of the adult learner, but rather the result of a "mapping problem" or computational difficulty during online processing of L*n*. Much reputable research (see: Alarcón, 2011 and Montrul et al., 2008) concludes that traditional L2 learners, heritage learners, and adult native Spanish speakers all possess competence with grammatical gender, but grammatical gender knowledge might be stored, represented, and reproduced differently in the mental grammars of these different groups of speakers.

Furthermore, the assumption of a uniform and highly efficient native speaker is a view that has long been challenged in the L1 processing literature due to compelling evidence of individual variation in native speakers. Variability has been observed in both nonnative and native speaker processing of grammatical gender and differential behavior of masculine and feminine gender has been observed in that the masculine is commonly treated as a default form in both native and nonnative performance. Moreover, using morphological cues for gender assignment (i.e., following the prototypical pattern of gender marking on Spanish nouns) is not unique to adult learners and, along with treating the masculine gender class as default, appears to constitute general linguistic processing strategies of all language users.

With past research and theory in mind, this study aims to examine how linguistic variables, task demands, and individual factors affect performance (measured in both mean accuracy and intra-speaker variance) in both native speakers and adult/late learners. Furthermore, this study explores how being multilingual may impact linguistic performance and examines the nature of this multilingual effect as either typological or more general. In order to address these objectives, 115 participants, including 25 native speakers and 90 nonnative Spanish learners of varying proficiency levels were recruited and completed a Language Learner Profile Questionnaire, a Spanish proficiency test, four experimental tasks with varying task conditions, and a metalinguistic awareness exit survey. In what follows, the research questions and corresponding predictions grounded in previous research and theory will be presented.

2.8 Research Questions and Hypotheses

Informed by previous research and theory on differentiated performance in native and nonnative speakers, the role of implicit and explicit knowledge stored in procedural and declarative memory systems, the effect of task demands, research on the impact of individual differences and multilingualism, and with a complex systems approach that treats variation as central, this research is guided by four main questions and poses the following predictions accordingly:

QUESTION 1. How do native Spanish speakers and late learners of Spanish perform with grammatical gender agreement?

> **1.1** To what extent do the linguistic variables of noun gender class, morphological marking, domain of agreement, and target noun frequency modulate participant performance in both native speakers and adult learners?

HYPOTHESIS 1: It is predicted that native Spanish speakers will perform near or at ceiling with gender agreement and that adult learners will perform above chance level, but significantly below native speaker controls (Fernández-García, 1999; McCarthy, 2008; Montrul, Foote, & Perpiñán, 2008). The linguistic variables analyzed will impact performance of both native and nonnative speakers (Beatty-Martínez & Dussias, 2019; Eddington & Hualde, 2008; Eddington, 2002), but Spanish learners will be significantly more affected and will exhibit higher accuracy with overtly-marked, masculine nouns (Black & Tararova, 2020; Foote, 2015; Gamboa, 2012; Montrul et al., 2008), whereas such linguistic factors may be detected for native speakers too, but will likely not reach significance as their effect size will be smaller. Furthermore, it is expected that learners will be more accurate with N-Adj. agreement than with Det.-N agreement (Black & Tararova, 2020; Grüter et al., 2012). It is predicted that among late Spanish learners, proficiency will interact with the relative effect of the linguistic variables such that as proficiency increases, the effect of the linguistic variables analyzed will decrease since at higher levels of proficiency, there will be less variation in scores to significantly interact with other variables. Likewise, it is predicted that

Spanish learners will show significantly higher inter- and intra-speaker variation since, as a group, late learners tend to be more impacted than native speakers by individual differences (e.g. Andringa & Dąbrowska, 2019) and because late learners tend to depend more extensively on explicit knowledge, which is subject to greater variability at the individual level (R. Ellis, 2005) and is stored in declarative memory, which is less systematic (Paradis, 2004). Much less inter- and intra-speaker variation is expected among native Spanish speakers since the target structure analyzed is largely part of the core grammar of the language and therefore falls within the domain of Basic Language Cognition, which is known to be shared by all native speakers (Hulstijn, 2015) and is part of language competence, which is largely implicit and systematic (N.C. Ellis, 1996, 2015; R. Ellis, 2005); furthermore, in native speakers, grammar is posited to be stored in procedural long term memory, which is known to be more systematic and therefore subject to far less variation (Paradis, 2004; Ullman 2001, 2004).

QUESTION 2. How do the task demands (time constraint, input modality) impact accuracy and variation in performance in both native speakers and adult learners?

HYPOTHESIS 2: Task effects are predicted for both types of task demands: time constraint and input modality. With regards to time constraint, lower accuracy scores and higher variation are expected for adult learners on Tasks 2-4 due to the imposed time constraint and focus on meaning, thus placing increased demands on working memory thereby requiring more automatic processing (R. Ellis, 2005) which, in turn, draws on implicit linguistic competence, posited to be present to a lesser extent in nonnative speakers, particularly at lower proficiency levels (N.C. Ellis, 2015). Furthermore, with regards to a dynamic systems approach to language development, stages of system flux, characterized by high levels of variation (Larsen-Freeman, 2012; N.C. Ellis & Larsen-Freeman, 2006), are expected in a developing learner grammar, whereas more stable, attractor states displaying less variation are expected for more consolidated mental grammars at higher levels of competence, such as in highly proficient learners and native speakers. With regards to stimuli presentation modality, lower accuracy and higher variation are expected for the aural modality due to the increased demands of auditory processing (Murphy, 1997). Native speakers are expected to show some effect of stimuli modality, but to a far lesser extent than nonnative speakers (Murphy, 1997).

QUESTION 3. To what extent can individual differences between learners' proficiency (self-report & measured), metacognitive awareness level, metalinguistic awareness of the target structure, motivation type and level, attitudes about the target language and target language community, Spanish language use, and linguistic repertoire/experience (presence/absence of grammatical gender) predict or account for accuracy and variation in performance?

> **3.1** To what extent can these individual differences account for differentiated performance in speeded vs. self-paced tasks (time constraint) and auditory vs. written tasks (stimuli modality)?

HYPOTHESIS 3: Due to the fact that individual learner differences have a large determining influence on target language achievement and performance in late learners (e.g., Andringa & Dąbrowska, 2019), it is expected that all individual variables analyzed will have some effect on performance, albeit with varying effect sizes. Although native Spanish speakers will show some degree of variation in these individual variables, the variables will likely not have a significant effect on performance. For late learners, proficiency is expected to have the largest impact on performance (Danesh & Shahnaazari, 2020; Hulstijn, 2019; Leaver, 2003). Furthermore, some significant interaction among variables is expected, particularly between metacognitive awareness and metalinguistic awareness (i.e., students who are more aware of their learning process and associated strategies may be more likely to notice linguistic forms and patterns as part of those strategies), between metalinguistic awareness and linguistic repertoire/multilingualism (i.e., the more languages one speaks, the more likely they are to have more developed explicit awareness of language structure more broadly), and between motivation and attitudes (i.e., learners with more positive attitudes toward the language and the target language community are more likely to experience higher levels of motivation in the idealized L2 self). With regards to what extent individual differences will interact with the time constraint task factor, a negative association with proficiency is expected such that as proficiency increases, the effect of time constraint will decrease as highly proficient speakers have more extensively developed implicit linguistic competence that is more readily accessible under a time pressure (R. Ellis, 2005). Metacognitive awareness is predicted to be more associated with untimed, self-paced tasks that thereby allow for access to explicit linguistic knowledge, but crucially not predictive of performance on timed tasks that require greater use of implicit linguistic knowledge (Black & Tararova, 2020). No other individual learner differences are expected to interact with the linguistic variable effects.

QUESTION 4. Is there a multilingual effect? Do participants who report experience with more languages perform differently from those who report experience with only English and Spanish?

> **4.1** To what extent can the presence of grammatical gender (binary or otherwise) in adult learners' prior linguistic repertoire account for accuracy in their performance or is an observed multilingual effect largely independent of typological similarity? Does the number of additional languages reported have an effect on performance?

HYPOTHESIS 4: Based on previous research demonstrating a positive effect of multilingualism on novel language learning (e.g., Black & Tararova, under review; Antoniou et al., 2015; Tremblay & Sabourin, 2012; Kaushanskaya & Marian, 2009; Klein, 1995), it is predicted that learners who report at least some proficiency with an additional language will exhibit higher accuracy than exclusively bilingual learners. Furthermore, this effect is hypothesized to be greater for learners who know more than three languages (Dmitrenko, 2017; Thompson, 2013; Herdina & Jessner, 2002). Some research has indicated that the multilingual advantage is somewhat independent of the typological similarity between languages in the learner's linguistic repertoire and therefore is more associated with general gains in the cognitive-linguistic system than mere transfer effects between typologically similar languages (e.g. Black & Tararova, 2020; Hirosh & Degani, 2018; Kaushanskaya & Marian, 2009). Therefore, it is predicted that having competence in an additional language that also has grammatical gender will not have a significant effect that is independent of the general effect of multilingualism; in other words, among the subgroup of multilingual participants, those with and without grammatical gender already instantiated in their prior linguistic repertoire are predicted to display no statistically significant differences in performance. Multilingualism is not expected to have an effect on native speaker performance since native speakers will likely perform at or near ceiling and therefore there is limited variation by which to observe the impact of other explanatory variables.

In the following chapter, the experimental methodology will be described in detail, including a description of the study participants, the procedure that was followed, the materials and tasks that were utilized, as well as a description of the study variables examined and how the resulting data was organized and analyzed.

Chapter 3: Methodology

3.1 Participants

Participants (N = 115) in this study consisted of both adult (+18 years of age) native Spanish speakers (n = 25) and adult nonnative instructed learners (n = 90). The nonnative/learner group consisted of late beginner and intermediate Spanish students recruited from first- and second-year Spanish courses at a large English-speaking Canadian university. In addition, advanced learners were recruited from the third- and fourth-year Spanish courses. Spanish did not necessarily need to be participants' L2 in order of acquisition nor did it need to be their only other nonnative language known, as this study endeavors to take a more linguistically diverse and multilingual perspective of learner performance, particularly given the rich linguistic context exhibited in higher education in Canada. Native Spanish speakers were recruited from a graduate program in Hispanic Studies at the same Canadian university and by word of mouth and friend of friend methods. The only qualifying criterion to participate as a native Spanish speaker was that Spanish must be the participant's *only first* language from birth; however, native speakers may have proficiency in other languages as well (i.e., via sequential bi-/multilingualism). In particular, English and French were expected as additional languages among the native Spanish speakers recruited in Canada. In order to work with human participants, this project received ethics approval via the Non-Medical Research Ethics Board (NMREB), under the Project ID: 11828.

Participants: incidence of multilingualism

This study endeavors to take a more multilingual perspective by recruiting a variety of participants from diverse language backgrounds. The Language Learner Profile Questionnaire, as will be described in the Materials and tasks section to follow, collected a variety of personal data from participants, including if and what additional languages (beyond English and Spanish) participants knew and at what proficiency level. Data collected regarding multilingualism was therefore exclusively based on self-report, and participants' proficiency in any of their reported languages (other than Spanish) was not tested. Figure 1 demonstrates the portion of additional languages reported by all multilingual participants (n=86), including both native Spanish speakers and adult learners of Spanish.

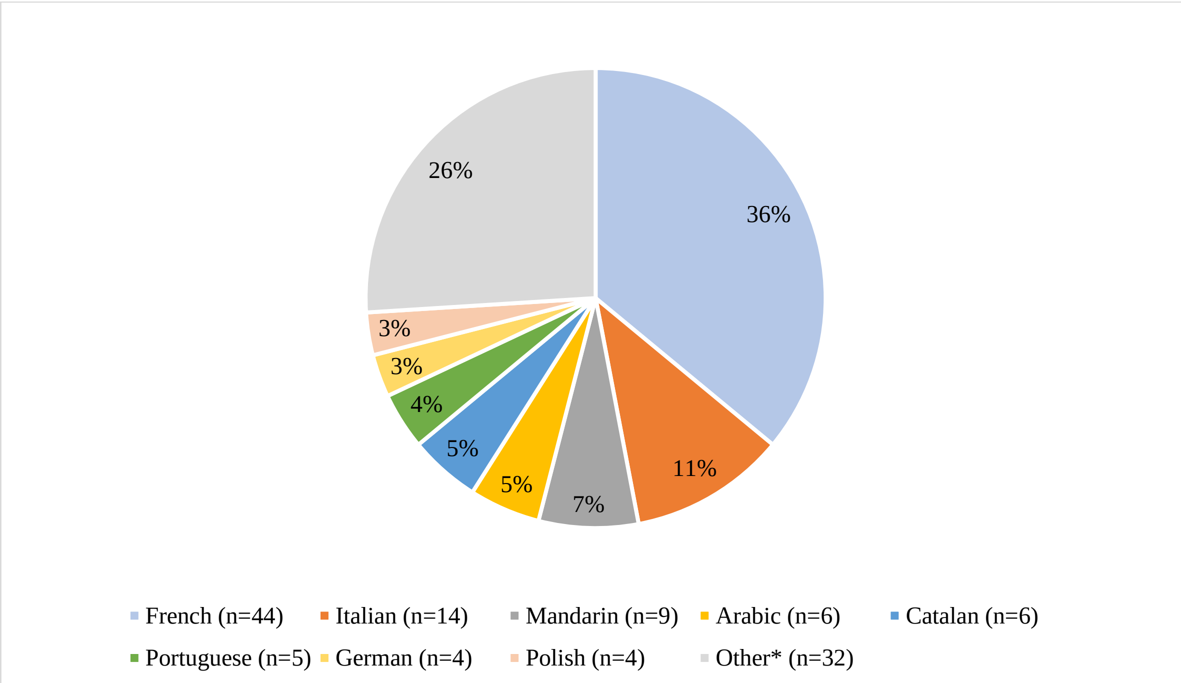


Figure 1. Additional languages reported beyond English and Spanish by all multilingual participants. Percentages shown represent the proportion of the multilingual participant sample only (n=86). *See Table 5 for a complete list of all additional languages reported.

Data was also collected on the order of acquisition of the additional languages reported by the multilingual participants as either *first*, *second*, or *third* additional language beyond English and Spanish. This information is summarized in Table 5, including the total count of participants who reported knowledge of each additional language, the percent of the total sample, and the total count of the number of multilingual participants who reported each of the additional languages as either a first, second, or third additional language beyond Spanish and English. In total, 31 additional languages were reported by multilingual participants, of which French (n=44), Italian (n=14), and Mandarin (n=9) were the most common.

Table 5. Additional languages reported (beyond English and Spanish) by all multilingual participants according to their order of acquisition after Spanish (L1/L2) and English (L1/L2).

			order of acquisition (after Spanish & English)		
Additional Language Reported	**Total Count**	**% Total**	***first* additional language**	***second* additional language**	***third* additional language**
French	44	35.48%	27	16	1
Italian	14	11.29%	9	1	4
Mandarin	9	7.26%	8	1	0
Arabic	6	4.84%	5	1	0
Catalan	6	4.84%	4	2	0
Portuguese	5	4.03%	2	2	1
German	4	3.23%	2	2	0
Russian	3	2.42%	3	0	0
Romanian	3	2.42%	3	0	0
Polish	4	3.23%	2	2	0
Ukrainian	2	1.61%	1	1	0
Korean	2	1.61%	2	0	0
Gujarati	2	1.61%	2	0	0
Japanese	2	1.61%	0	1	1
Tamil	2	1.61%	2	0	0
Vietnamese	1	0.81%	1	0	0
Malayalam	1	0.81%	1	0	0
Farsi	1	0.81%	1	0	0
Tagalog	1	0.81%	1	0	0
Swahili	1	0.81%	1	0	0
Thai	1	0.81%	1	0	0
Armenian	1	0.81%	1	0	0
Cantonese	1	0.81%	1	0	0
Sinhalese	1	0.81%	1	0	0
Haitian Creole	1	0.81%	1	0	0
Yoruba	1	0.81%	1	0	0
Bahamian Creole	1	0.81%	1	0	0
Danish	1	0.81%	1	0	0
Indonesian	1	0.81%	1	0	0
Hungarian	1	0.81%	0	1	0
Kinyarwanda	1	0.81%	0	1	0

As Figure 2 shows, the majority of adult Spanish learners self-reported as "multilingual" (n=74), comprising approximately 82% of the learner sample, whereas only 18% of the learners (n=16) reported only knowledge of English and Spanish (i.e., "bilingual"). Native Spanish speakers were nearly evenly split between "bilinguals" (n=13), reporting knowledge of Spanish as L1 and English as L2, and "multilinguals" (n=12), reporting knowledge of additional languages beyond Spanish (L1) and English (L2).

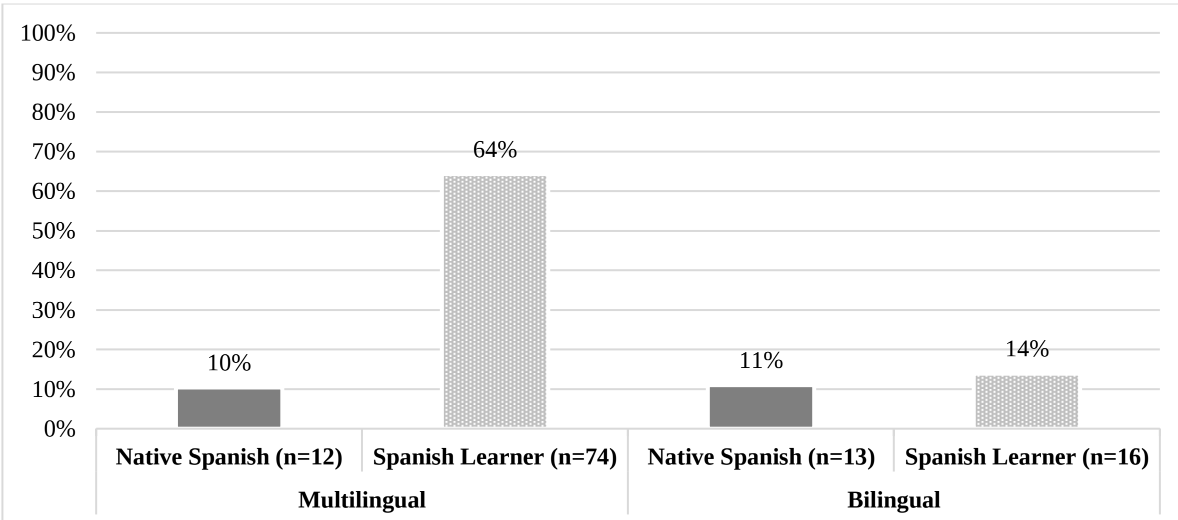


Figure 2. Frequencies of multilingualism self-reported per participant group on the Language Learner Profile Questionnaire. Percentages shown represent the proportion of the entire participant sample (N=115).

In addition to the presence of multilingualism in the sample, multilingual participants' prior linguistic repertoire was taken into account and their additional languages reported were classified according to whether they also exhibited a grammatical gender system. As Figure 3 demonstrates, a grammatical gender system was already present in the majority of multilingual learners' prior linguistic repertoires (n=56), although some multilingual learners reported knowledge of languages that do not feature a grammatical gender system (n=18). It is important to note that all native Spanish speakers, whether multilingual or not, had grammatical gender present in their respective linguistic repertoires by virtue of being native speakers of Spanish, a gendered language.

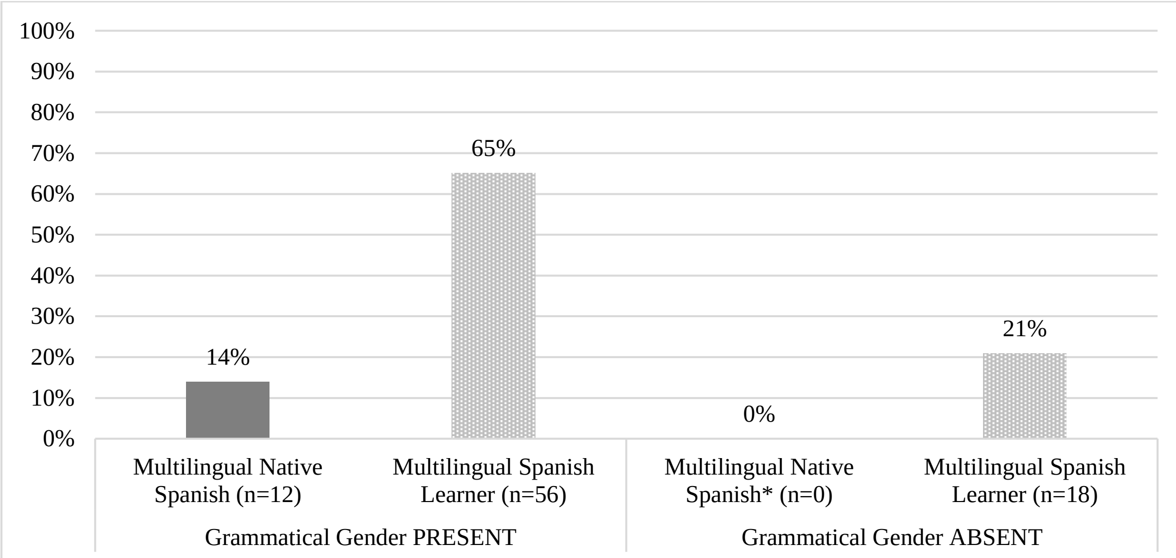


Figure 3. Frequency of the presence of grammatical gender in the prior linguistic repertoire of multilingual participants, organized per participant group. Percentages shown represent the proportion of the multilingual participant sample only (n=86). ****Note:*** *all native Spanish speakers have grammatical gender present in their linguistic repertoires by virtue of being native speakers of Spanish, a gendered language.*

Furthermore, multilingual learners with grammatical gender already instantiated in their prior linguistic repertoire were further divided according to the grammatical gender system type as binary (i.e., masculine/feminine) or non-binary (i.e., masculine/feminine/neuter or other). In the vast majority of cases, the additional gendered languages reported were also of the binary subtype (n=52), whereas very few multilingual learner participants' additional gendered languages reported were of a non-binary gender subtype (n=4), as shown in Figure 4. Once again, it is important to note that all native Spanish speakers, whether multilingual or not, had binary grammatical gender present in their respective linguistic repertoires by virtue of being native speakers of Spanish, a binary gendered language.

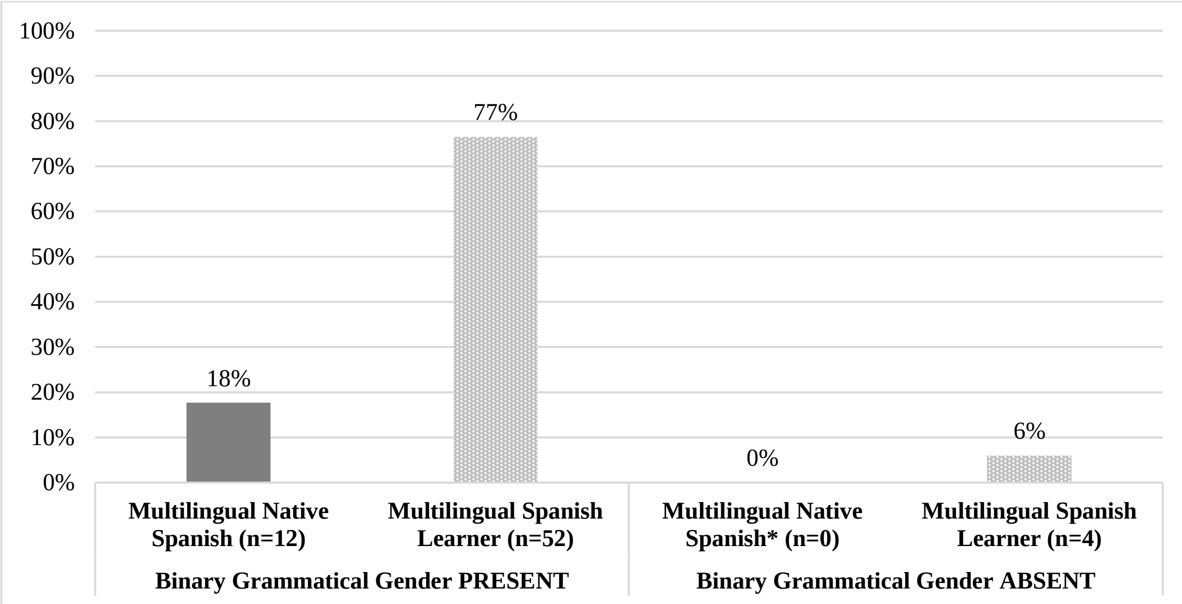


Figure 4. Frequency of the presence of *binary* grammatical gender in the prior linguistic repertoire of multilingual participants, organized per participant group. Percentages shown represent the proportion of the multilingual participant sample with grammatical gender present in their prior linguistic repertoire (n=68). ****Note:** all native Spanish speakers have binary grammatical gender present in their linguistic repertoires by virtue of being native speakers of Spanish, a binary gendered language.*

Finally, data was collected from multilingual participants regarding the number of additional languages reported (beyond English and Spanish) and a summary of this data is provided in Figure 5. Among multilingual learners, knowledge of one additional language was the most commonly reported (n=44) in the sample, followed by two additional languages (n=24), and finally a few multilingual learner participants also reported knowledge of three additional languages (n=6). For multilingual native Spanish speakers, the vast majority reported just one additional language (n=10), while two (n=1) and three (n=1) additional languages were reported by just one multilingual native speaker each, as shown in Figure 5.

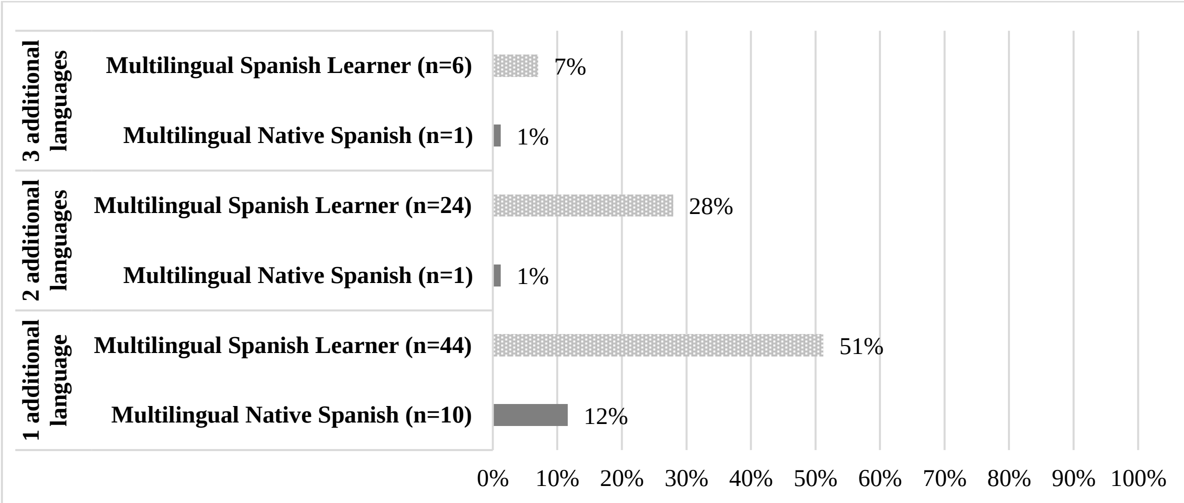


Figure 5. Number of additional languages reported (beyond English and Spanish) by multilingual participants, organized by participant group. Percentages shown represent the proportion of the multilingual participant sample only (n=86).

3.2 Procedure

During all experimental tasks, participants were audio and video recorded via Zoom. The pre-experimental task, consisting of the Letter of Information & Consent form and the Language Learner Profile Questionnaire, was completed prior to the recorded Zoom session via Qualtrics (an online survey platform licensed by the University). Participants received the Qualtrics link via email when they expressed interest in participating in the study and were also given an assigned participant code at this time which they entered on the Qualtrics form that allowed their Qualtrics responses to be linked to their Zoom session. The pre-experimental task took approximately 20 minutes in total to complete online prior to the scheduled Zoom session. All tasks (except for the pre-experimental task in Qualtrics) were completed during an individual synchronous Zoom session with the researcher and presented in PowerPoint slides through the "share screen" function in Zoom. The Spanish proficiency test (Task 1) was the first task administered, followed by experimental Tasks 2-5, and finally the metalinguistic awareness exit survey (Task 6) was the last task participants completed, as detailed below. Participants were compensated in the form of an electronic Amazon.ca gift card in the amount of $20 CAD delivered via email after completing the Zoom session. In addition, participants who were taking a Spanish course at the time of the study received 1% course credit for their participation in research. All tasks in total took approximately 50-60 minutes to complete for learner participants and approximately 40-45 minutes for native Spanish speakers, via a synchronous individual Zoom session.

3.3 Materials and tasks

3.3.1 Pre-experimental task: LOI & Consent + Language Learner Profile Questionnaire

The letter of information (LOI) and the consent form were integrated into the Qualtrics survey and were the first thing participants saw when they clicked on the Qualtrics link. The Language Learner Profile Questionnaire included sections on self-reported proficiency, language learning history (i.e., linguistic repertoire) and use, attitudes, motivation, and metacognitive awareness. The sections that compose the questionnaire are detailed below in Table 6. See *Appendix A* for the full questionnaire.

Table 6. Components of the Language Learner Profile Questionnaire.

Questionnaire component	**Description**
Linguistic profile:	Comprises language history, use, self-reported proficiency, and language attitudes, adapted from the Bilingual Language Profile Questionnaire (BLP: Birdsong, Gertken, & Amengual, 2012).
Motivation:	The statements utilized in the motivation section of the questionnaire were adapted to Spanish from Aubrey's (2014) motivation questionnaire, which was constructed based on the L2 Motivational Self System framework (Dörnyei, 2005, 2009), and was originally adapted from Taguchi et al. (2009) and subsequently validated (Aubrey & Nowlan, 2013). The self-report motivation questionnaire utilizes statements to be ranked by participants along a 5-point Likert scale in the domains of motivated learning behavior, ideal L2 self, ought-to L2 self, and the L2 learning experience, as delineated by Dörnyei (2005).
Metacognitive awareness:	The statements to evaluate level of metacognitive awareness were taken from the Metacognitive Awareness Inventory for Language Learning (MAILL) questionnaire (Black & Tararova, 2020), which was shortened and adapted to Spanish from the original Metacognitive Awareness Inventory (MAI) by Schraw & Dennison (1994). These statements concern knowledge about cognition and regulation of cognition in the context of learning Spanish in which participants are instructed to self-report how often the statements apply to them along a 5-point Likert scale.

The Language Learner Profile Questionnaire consisted of eight different sections: general personal information, information about the first language of the participant and the first languages of the participant's parents, language history, language use, self-reported language proficiency in all languages known, language attitudes about English, Spanish, and two other (optional) reported languages, language learner awareness (metacognitive awareness), and motivational orientation.

Section A: Personal Information featured a total of four questions about gender identity, age, occupation, and educational attainment level. *Section B: First Language* included six questions about the participant's native language, their mother's and father's native language(s), what language(s) they spoke at home as a child, and whether they are currently most comfortable with their first language. Figure 6 demonstrates the linguistic diversity observed in the present sample in terms of participants' first languages (L1) reported. In total, there were 21 different first languages reported, of which the most common were English (n=55; 48% of the sample) and Spanish (n=25; 22% of the sample).

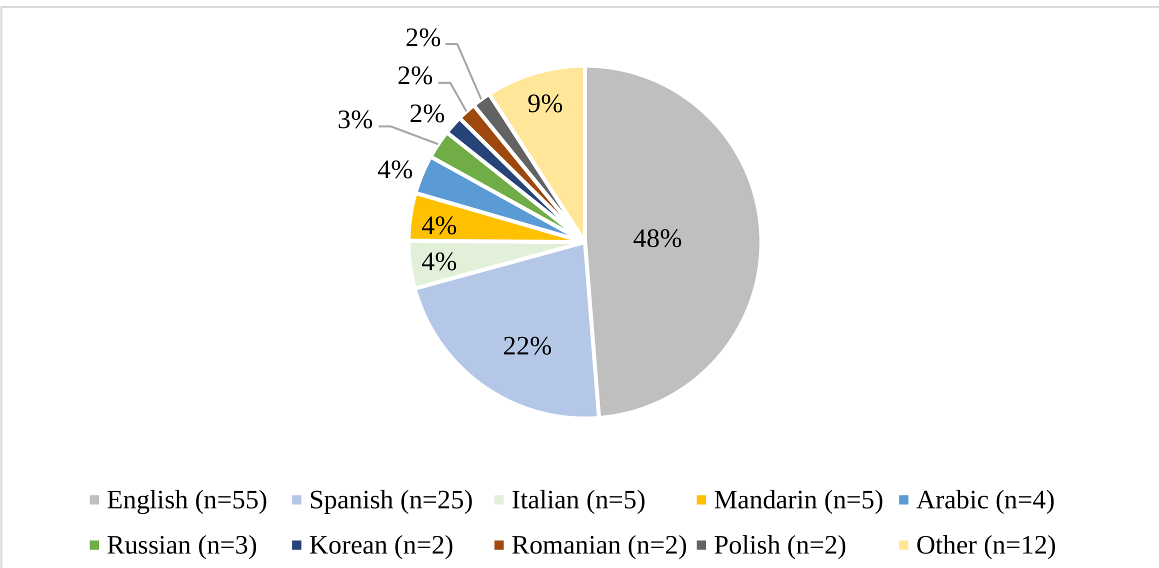


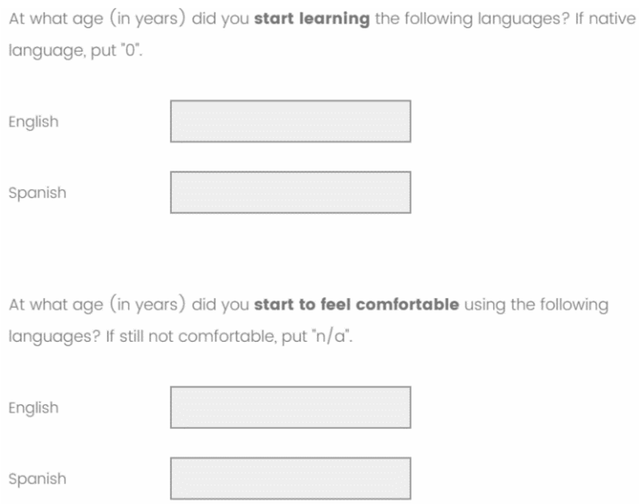
Figure 6. Proportion of **first languages (L1)** reported by all participants on the Language Learner Profile Questionnaire. The "Other" category includes one participant each from the following languages: Chinese, Ukrainian, Hungarian, Tagalog, Armenian, Sinhalese, Creole, Yoruba, French, Indonesian, Gujarati, German.

Section C: Language History included six questions about their age of acquisition (AoA)[1] of English and Spanish, the age at which they started to feel comfortable speaking each language,

[1]AoA data was not analyzed for the current study as all nonnative learners of Spanish were late/adult learners (post-puberty, after the age of 13 years) in order to qualify to participate. Furthermore, previous research indicates that

how many years of formal coursework in each language, and years spent in a country, family, and work environment where English and Spanish are spoken. See Example 1 below for a sample of the language history questions.

Example 1. Sample questions from ***Section C: Language History*** of the Language Learner Profile Questionnaire.



At what age (in years) did you **start learning** the following languages? If native language, put "0".

English

Spanish

At what age (in years) did you **start to feel comfortable** using the following languages? If still not comfortable, put "n/a".

English

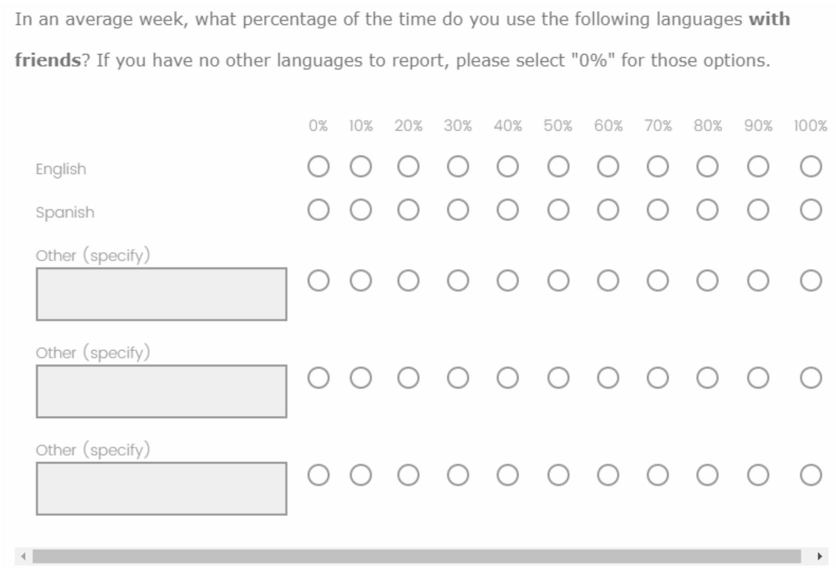
Spanish

Spanish language history scores were calculated by taking the average of each participant's age of onset (AO) score and Spanish experience score and then dividing that figure by the participant's age at the time of the study. Then participant group averages were calculated for native Spanish speakers and adult Spanish learners. As demonstrated in Figure 7, native Spanish speakers exhibited, as expected, a much higher Spanish history score (M = .79, SD = 0.12) than their adult learner peers (M = .12, SD = 0.12), as it was very difficult for the (late/adult) learner participants to compensate for not having an early age of onset of Spanish, despite, in some cases, having years of coursework and experience in a country and work environment where the language is spoken.

Section D: Language Use included five questions that asked participants to quantify how often they use each of their reported languages as a percentage in a typical week with friends, family, at school/work, when engaging in self-talk, and when counting. See Example 2 below for a sample of the language use questions.

AoA is no longer a significant predictive factor in late language acquisition (see: Birdsong, 2005, 2006; Birdsong & Vanhove, 2016; Singleton & Muñoz, 2011).

Example 2. Sample question from ***Section D: Language Use*** of the Language Learner Profile Questionnaire.


In an average week, what percentage of the time do you use the following languages **with friends**? If you have no other languages to report, please select "0%" for those options.

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
English											
Spanish											
Other (specify)											
Other (specify)											
Other (specify)											

Spanish use scores were calculated for each participant by determining the average percentage reported of Spanish language use across all five contexts: with friends, with family, at school/work, in self-talk, and when counting. Then participant group averages were calculated for native Spanish speakers and adult Spanish learners. Native Spanish speakers used Spanish overall much more often in a typical week (M = .56, SD = 0.21) than their learner peers (M = .10, SD = 0.08), as shown in Figure 7.

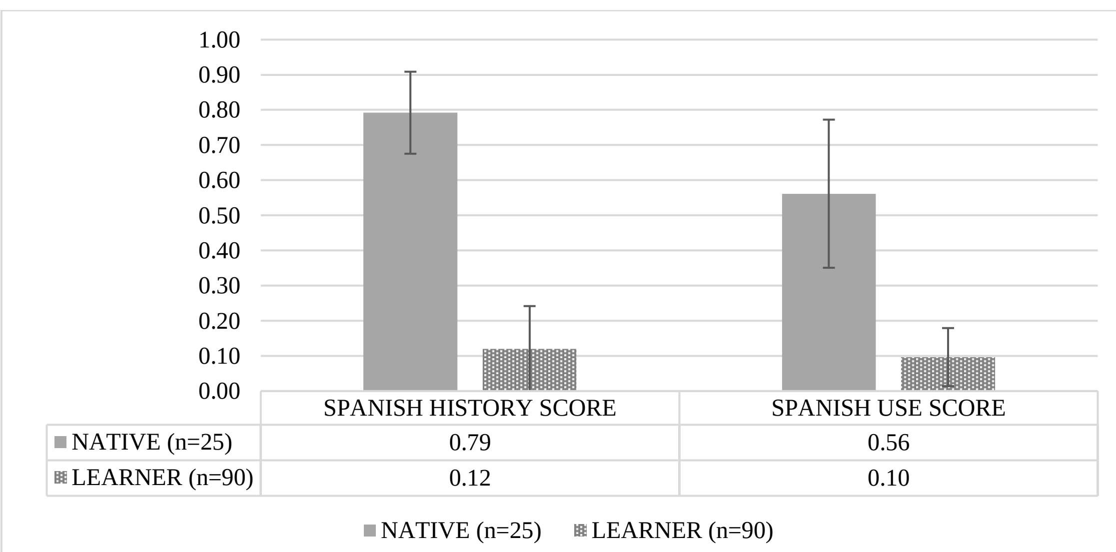
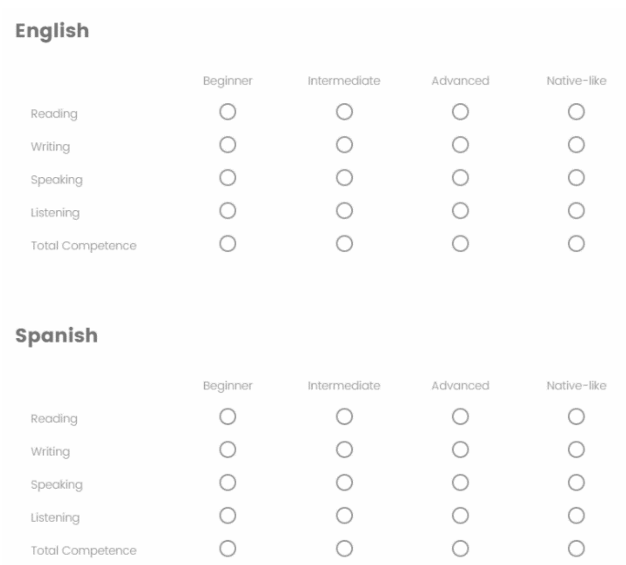


Figure 7. Average **Spanish language history** and weekly **Spanish language use** scores as self-reported by each participant group on the Language Learner Profile Questionnaire. Standard deviation (SD) bars are provided.

Section E: Language Proficiency: Self-assessment consisted of separate sections for each language in which participants were asked to rate their current proficiency level on a scale of “beginner”, “intermediate”, “advanced”, or “native-like” in the domains of reading, writing, speaking, listening, and total competence. See Example 3 below for some sample prompts of the self-assessed proficiency portion of the language learner profile questionnaire.

Example 3. Sample prompts from ***Section E: Language Proficiency: Self-assessment*** of the Language Learner Profile Questionnaire.



English

	Beginner	Intermediate	Advanced	Native-like
Reading	○	○	○	○
Writing	○	○	○	○
Speaking	○	○	○	○
Listening	○	○	○	○
Total Competence	○	○	○	○

Spanish

	Beginner	Intermediate	Advanced	Native-like
Reading	○	○	○	○
Writing	○	○	○	○
Speaking	○	○	○	○
Listening	○	○	○	○
Total Competence	○	○	○	○

Self-reported Spanish proficiency was quantified in the following manner: responses of “beginner” were converted to a score of 0.25; “intermediate” was converted to a score of 0.50; “advanced” was converted to a score of 0.75; and “native-like” was converted to a score of 1.00. Average self-reported Spanish proficiency was calculated by computing the average self-reported level on each of the subsections: reading, writing, speaking, listening, and total competence. Participants’ self-reported Spanish proficiency scores are reported together with the tested Spanish proficiency scores in the description of the Task 1: Spanish Proficiency Test below.

Section F: Language Attitudes asked participants to think about how the languages they speak reflect their identity, values, and how they want to be perceived in society and prompted participants to respond to four statements along a scale of “strongly disagree” to “strongly agree”.

The objective was to evaluate to what extent participants identified with the target language community and culture. See Example 4 below for the four statements to which participants were asked to respond.

Example 4. Statement prompts for Spanish from ***Section F: Language Attitudes*** of the Language Learner Profile Questionnaire. Participants also responded to the same statements for each of the known languages they reported.


Spanish

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
I **feel like myself** when I speak **Spanish**.	○	○	○	○	○
I **identify with** a **Spanish-speaking culture**.	○	○	○	○	○
It is **important to me** to **use Spanish like a native speaker**.	○	○	○	○	○
I **want others to think** that I'm a **native speaker of Spanish**.	○	○	○	○	○

Spanish attitude responses were quantified such that “strongly disagree” was converted to -2; “somewhat disagree” was converted to -1; “neither agree nor disagree” was converted to 0; “somewhat agree” was converted to +1; “strongly agree” was converted to +2. Average Spanish attitude scores were computed for each participant by calculating the average response on all four prompts. Native Spanish speakers exhibited entirely positive attitude scores (M = 1.66, SD = 0.50) toward their native language community whereas adult learners were more varied in their attitudes (M = 0.11, SD = 1.05), exhibiting almost equally positive and negative attitude scores, as shown in Figure 8.

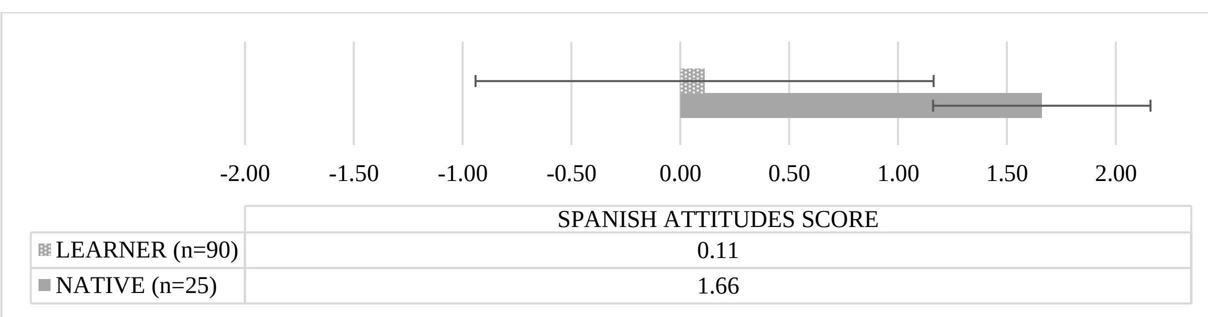


Figure 8. Average scores on **attitudes** about the Spanish language and target language community as self-reported by each participant group on the Language Learner Profile Questionnaire. Standard deviation (SD) bars are provided.

Only learners of Spanish completed the last two sections of the Language Learner Profile Questionnaire. *Section G: Language Learner Awareness* was designed to assess learner's metacognitive awareness level, that is, to what extent they are self-aware of themselves as language learners and how they employ certain strategies to learn the language. There was a total of twenty statements to which participants were asked to respond on a scale of "never" to "always" based on the frequency with which the statement applies to them in a typical week during the past month. See Example 5 below for some sample statements included in this section.

Example 5. Statement prompts from ***Section G: Language Learner Awareness*** of the Language Learner Profile Questionnaire.

	Never	Sometimes	About half the time	Most of the time	Always
I know what information is most important to learn.	○	○	○	○	○
I create my own examples in Spanish to make new information more meaningful.	○	○	○	○	○
When I don't understand something in Spanish, I ask others for help (teacher, classmates, tutors, etc.).	○	○	○	○	○
I focus on overall meaning rather than specifics.	○	○	○	○	○

Responses on the language learner awareness section were quantified in the following manner: "never" was converted to a score of 0; "sometimes" was converted to a score of 1; "about half the time" was converted to a score of 2; "most of the time" was converted to a score of 3; and "always" was converted to a score of 4. Two groups were subsequently formed: "more aware" learners (average scores of '3' or more) and "less aware" learners (average scores of less than '3'). As illustrated in Figure 9, most learners were considered "less aware" (n=68) in their knowledge about and regulation of cognition, yet approximately 25% of the sample of late learners of Spanish were considered "more aware" (n=22) of themselves as language learners and the language learning process.

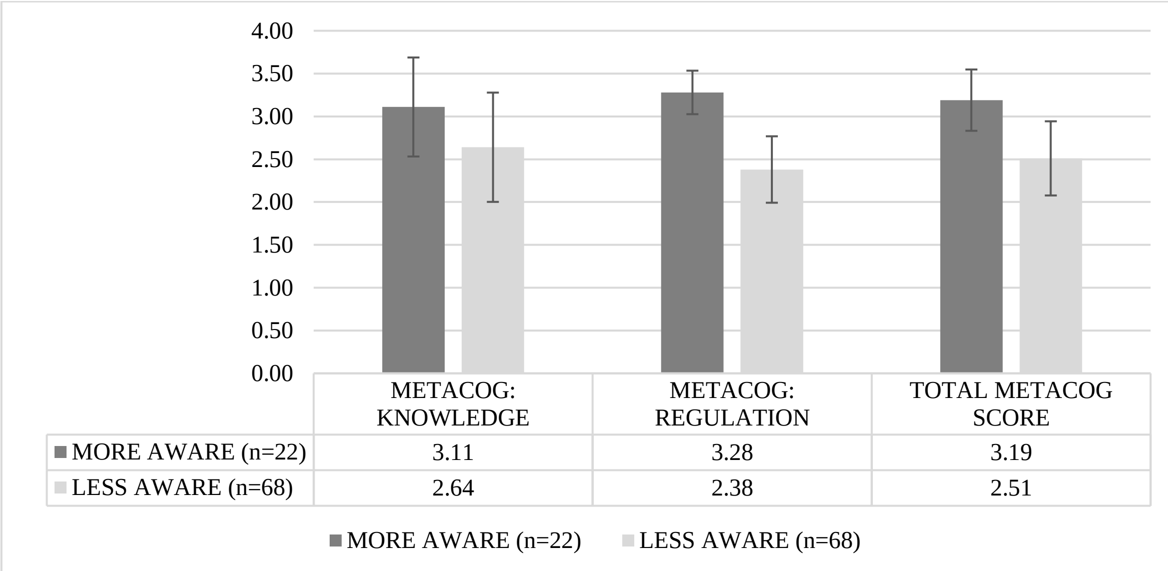


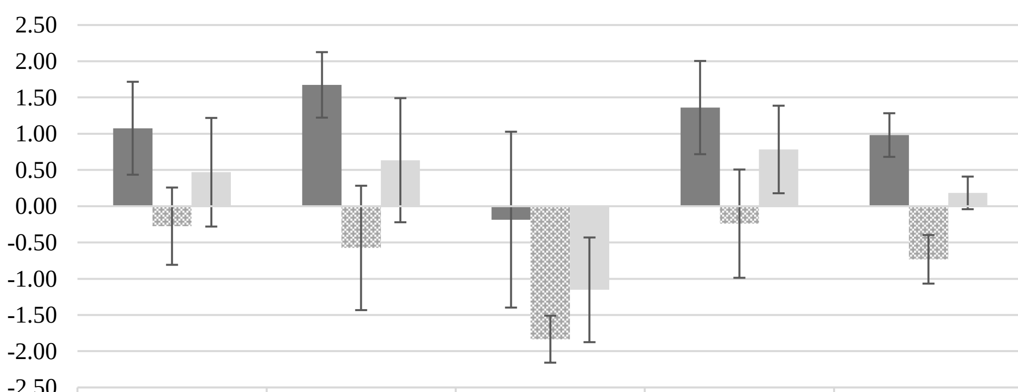
Figure 9. Average language **learner awareness** scores per component of metacognitive awareness and total as self-reported by each participant on the Language Learner Profile Questionnaire. Standard deviation (SD) bars are provided.

Section H: Motivational Orientation was the final section of the questionnaire and collected information from the adult learner participants on both the nature and degree of their individual motivation to learn Spanish and included four subsections: motivated learning behavior, ideal L2 self, ought-to L2 self, and motivated learning experience. There were five statements in each subsection to which participants were asked to respond along a scale from “totally disagree” to “totally agree”. See Example 6 below for some sample statements included in this section.

Example 6. Statement prompts from ***Section H: Motivational Orientation*** of the Language Learner Profile Questionnaire.

	Strongly disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Strongly agree
If a Spanish course was offered at university or somewhere else in the future, I would like to take it.	○	○	○	○	○
I think I am doing my best to learn Spanish.	○	○	○	○	○
I am prepared to expend a lot of effort in learning Spanish.	○	○	○	○	○
Compared to my classmates, I think I study Spanish relatively hard.	○	○	○	○	○

Reponses on the motivational orientation section were quantified in the following manner: “strongly disagree” was converted to a score of -2; “somewhat disagree” was converted to a score of -1; “neither agree nor disagree” was converted to a score of 0; “somewhat agree” was converted to a score of +1; and “strongly agree” was converted to a score of +2. Three groups were subsequently formed: “positive” motivational orientation (average scores of +1 or more), “negative” motivational orientation (average scores of -1 or less), and “neutral” motivational orientation (average scores between -1 and +1). As illustrated in Figure 10, most learners exhibited an overall “positive” motivational orientation (n=56), while approximately 27% exhibited a more “neutral” motivational orientation (n=24), and a minority of learner participants exhibited a “negative” motivational orientation (n=10) towards the task of learning Spanish. With regards to the different components of motivation, the “ought-to L2 self” demonstrated the most negative scores overall in terms of others’ expectations, while the “ideal L2 self” demonstrated the most positive scores overall, meaning that late learners of Spanish tend to report being more motivated by an ideal projection of themself into the future as someone who has learned the language.



	MOTIV: BEHAVIOR	MOTIV: IDEAL L2 SELF	MOTIV: OUGHT-TO L2 SELF	MOTIV: EXPERIENCE	MOTIV: TOTAL SCORE
POSITIVE (n=56)	1.08	1.67	-0.19	1.36	0.98
NEGATIVE (n=10)	-0.28	-0.58	-1.83	-0.24	-0.73
NEUTRAL (n=24)	0.47	0.64	-1.15	0.78	0.18

POSITIVE (n=56) NEGATIVE (n=10) NEUTRAL (n=24)

Figure 10. Average **motivational orientation** scores per component of motivation and total motivation as self-reported by each participant on the Language Learner Profile Questionnaire. Standard deviation (SD) bars are provided.

3.3.2 Task 1: Spanish Proficiency Test

The proficiency test consisted of a cloze reading (12 targets) to assess grammatical competence and fill-in-the-blank sentences (20 targets) to assess vocabulary, taken directly from the McGill Spanish proficiency test and shortened. This test was originally developed by Duffield and White (1999) and subsequently used in many other studies (e.g., Duffield et al., 2002; Montrul, 2002; Bruhn de Garavito & Valenzuela, 2008; Cuza et al., 2013). The proficiency test was the first task participants completed during the synchronous session and was administered in real-time in PowerPoint slides over Zoom with written prompts to which participants responded orally with their answers recorded. See *Appendix B* for the proficiency test stimuli.

Tested Spanish proficiency scores were calculated in the following manner: all prompts received either a score of '1' for *correct* or a score of '0' for *incorrect*. *Part 1: Vocabulary* contained a total of 20 prompts with four (4) multiple choice options each, producing a total possible score of 20 points; *Part 2: Grammar* contained a total of 12 prompts with three (3) multiple choice options each, producing a total possible score of 12 points. Scores on each section were converted to a percentage and then were averaged to calculate the total tested Spanish proficiency score for each participant.

Figure 11 summarizes both tested and self-reported Spanish proficiency scores in native speakers (n=25) and adult learner (n=90) participants. As described above, Section E (*Language Proficiency: Self-assessment*) of the Language Learner Profile Questionnaire prompted participants to self-evaluate their current proficiency level in Spanish at the time of testing along a scale of "beginner", "intermediate", "advanced", or "native-like" in the domains of reading, writing, speaking, listening, and total competence. Figure 11 presents the average self-reported proficiency scores per participant group. Overall, native speakers reported a higher proficiency level of 96% corresponding, as expected, to the "native-like" categorization. Meanwhile, the adult learner participants reported an average proficiency in Spanish of 50%, corresponding to "beginner/intermediate" categorization. Regarding tested Spanish proficiency, native speakers scored on average 94% on the vocabulary component of the proficiency test while learners obtained an average vocabulary score of 53%. On the grammar component of the Spanish proficiency test, native speakers scored on average 93% while learners scored 51%. Total

Spanish proficiency scores were calculated by taking the average of tested vocabulary and grammar scores. In terms of total tested Spanish proficiency, native speakers scored an average of 93% and learners scored an average of 52%. Overall, native Spanish speakers consistently demonstrated significantly higher proficiency scores than their adult learner counterparts, as expected. Nonetheless, the standard deviation (SD) bars shown in Figure 11 demonstrate consistently high variation in adult learners' Spanish proficiency scores, justifying the formation of distinct proficiency groups for the analysis of the effect of Spanish proficiency on adult learner performance.

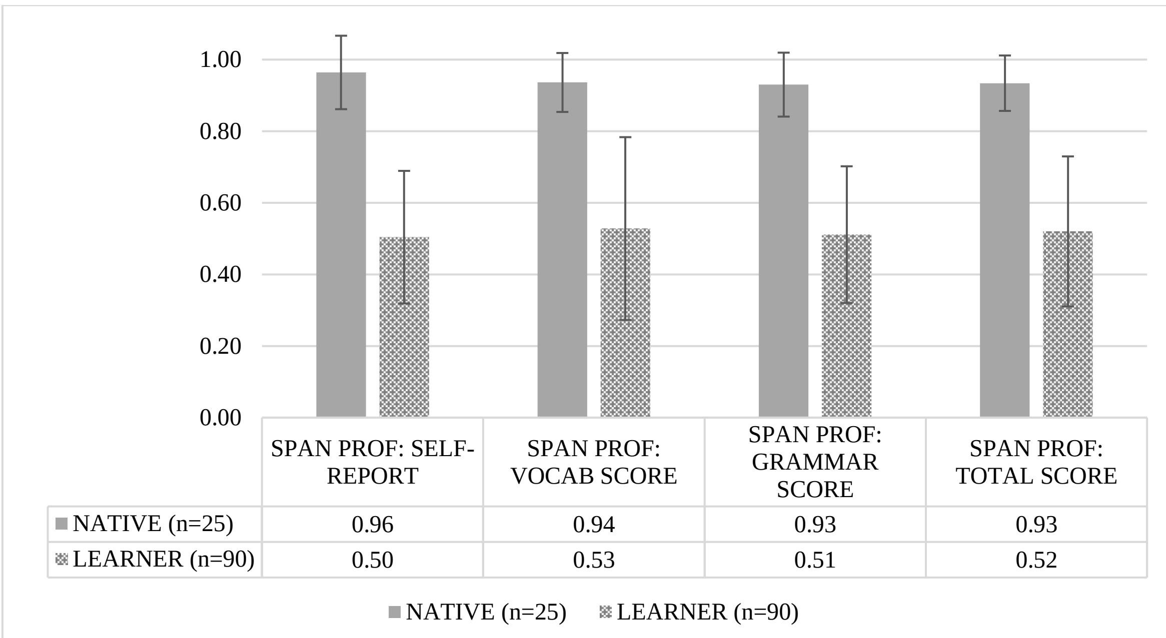
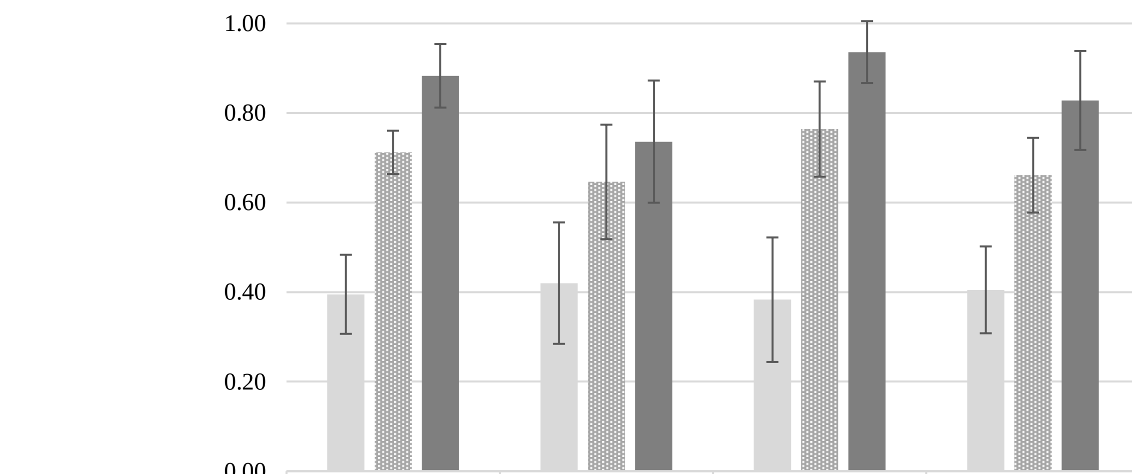


Figure 11. Tested and self-reported **Spanish proficiency** scores per participant group. Standard deviation (SD) bars are provided.

Figure 12 presents the self-reported and tested Spanish proficiency scores in adult learners only, divided into three proficiency groups: "beginner" (scores < .60), "intermediate" (.60 ≤ scores < .80), and "advanced" (scores ≥ .80). This categorization of learners into distinct proficiency groups revealed that nearly 70% of the learner sample (n=62) possessed a *beginner* level of proficiency at the time of testing and only roughly 15% of the learner participants tested at *intermediate* (n=14) and *advanced* (n=14) Spanish proficiency levels. During participant recruitment, the objective was to have more balanced proficiency groupings by recruiting from different course levels in the Spanish program; however, despite these efforts, the actual sample was certainly more skewed towards the beginner side of the proficiency spectrum and many participants self-reported higher proficiency levels than what they actually obtained via Spanish proficiency testing. Nonetheless, adult learner participants appear to be justifiably categorized into these three groupings, particularly when considering total tested Spanish proficiency scores (an average of grammar and vocabulary tested scores), as shown in Figure 12. Results of a one-way ANOVA confirm that the three proficiency groupings are, in fact, significantly different ($F(2, 29.7) = 309$, $p < .001$) when comparing their total tested Spanish proficiency scores.



	Total Tested Proficiency Score	Self-reported Proficiency Score	Tested Proficiency: Vocab Score	Tested Proficiency: Grammar Score
BEGINNER (n=62)	0.40	0.42	0.38	0.41
INTERMEDIATE (n=14)	0.71	0.65	0.76	0.66
ADVANCED (n=14)	0.88	0.74	0.94	0.83

Figure 12. Learner **self-reported and tested Spanish proficiency** scores compared per proficiency group. Standard deviation (SD) bars are provided.

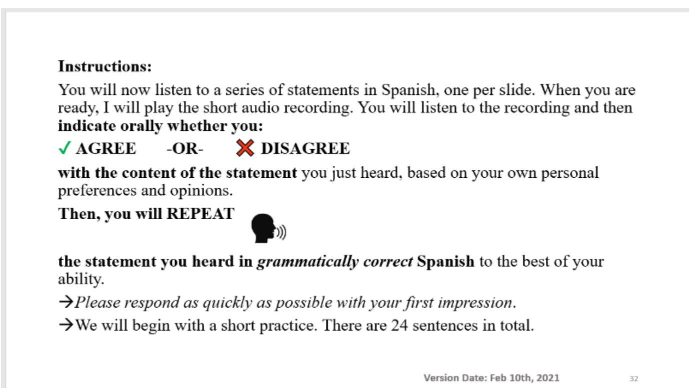
3.3.3 Experimental Tasks (4) + Exit Survey

The following experimental tasks to assess implicit and explicit linguistic competence of grammatical gender in Spanish were adapted from the battery of tests originally developed by R. Ellis (2005) and subsequently validated by other researchers among different learner populations (e.g., Bowles, 2011; Zhang, 2015). The experimental tasks detailed below were designed to create differentiated task demands such that tasks 3 and 4 were designed to examine the effect of modality of stimulus presentation in which task 3 featured stimuli presented in the auditory modality whereas task 4 featured stimuli presented in the orthographic/written modality. Similarly, tasks 2-4 conditioned the use of implicit linguistic knowledge (under a time constraint with a focus on meaning), whereas tasks 5 and 6 encouraged the use of explicit linguistic knowledge (by virtue of being self-paced with an explicit focus on form). Nonetheless, it should be noted that there can be no guarantee that 'task-as-workplan' directly corresponds to 'task-as-process' (e.g., Breen, 1989; Coughlan & Duff, 1994) given that learners are likely to draw on whatever resources they have at their disposal irrespective of which resources are the ones best suited to the task at hand. Therefore, we can only claim that these tasks have been designed to *predispose* participants to access one or the other knowledge types (implicit or explicit) in a more *probabilistic* than deterministic manner. The target nouns and adjectives used to test performance with grammatical gender in Spanish were all taken from the introductory course textbook, *Vistas* (6th edition), that was used at the time of the study in the Spanish program from which Spanish learners were recruited, in order to ensure learners' familiarity with relevant task vocabulary.

Task 2: Oral Imitation Task

This task consisted of 24 tokens total (12 grammatical; 12 ungrammatical) in the form of simple statements in Spanish that featured both gender agreement and agreement violations, presented aurally in randomized order. The grammatical tokens featured two types of agreement focus (6 Det-N agreement; 6 N-Adj. agreement) and the ungrammatical tokens featured two types of agreement errors (6 Det-N errors; 6 N-Adj. errors). So as not to cue gender assignment with the prompt itself, tokens with a focus or error on Det-N agreement were accompanied by an adjective invariable for gender (e.g., *verde*, "green"; *grande,* "big"; *importante,* "important", *interesante,* "interesting", etc.) and tokens with a focus or error on N-Adj. agreement contained either the invariable determiner *cada* ("each") or an invariable possessive determiner (e.g., mi, "my"; *tu*, "your"; *su*, "his/her"). Tokens were also evenly distributed by gender class (12 masculine; 12 feminine) and morphological marking (12 overtly marked/canonical; 12 non-overtly marked/noncanonical). One statement was presented per PowerPoint slide with a small speaker icon which the researcher pressed to play the audio recording. Participants were instructed to first listen to each statement (recorded by the same male native Spanish speaker) and then indicate orally agreement or disagreement with the provided statement based on their own personal preferences/opinions (e.g., ungrammatical, Det-N error: *la café es saludable*, "the-fem. coffeemasc. is healthy-invar.”; grammatical, N-Adj. focus: *mi casa es moderna*, “my-invar. house-fem. is modern-fem.”). First evaluating the auditory stimuli in order to express an opinion helped to focus participants’ attention on meaning rather than grammatical structure, which is considered essential for implicit processing of language (R. Ellis, 2005). After stating whether they agreed or disagreed with each statement, participants were then instructed to repeat the sentence they heard in *correct* Spanish. This second step assessed whether or not they were able to employ the adequate gender agreement operation to the statement they had previously processed for meaning. This task began with two practice sets featuring non-target stimuli (e.g., *me gusta correr todos los días*, “I like to run every day”; *quiero vivir en un edificio en el centro*, “I want to live in a building in downtown”). This task was also speeded in that participants were instructed to respond as quickly as possible with their first impression. See Example 1 below for the task instructions participants saw and an example prompt. See *Appendix C* for a full list of the stimuli used in this task.

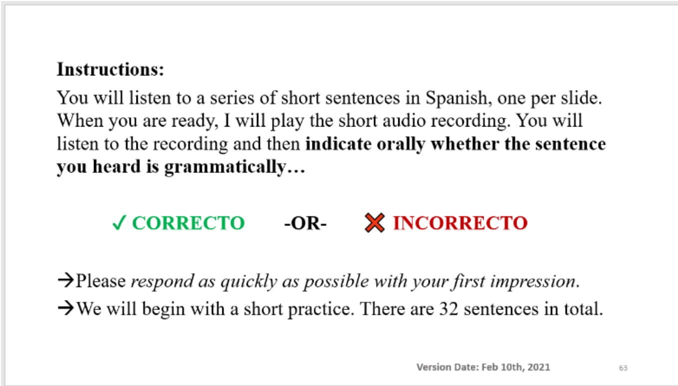
Example 1. Task 2 (Oral Imitation Task) instructions and sample prompt.

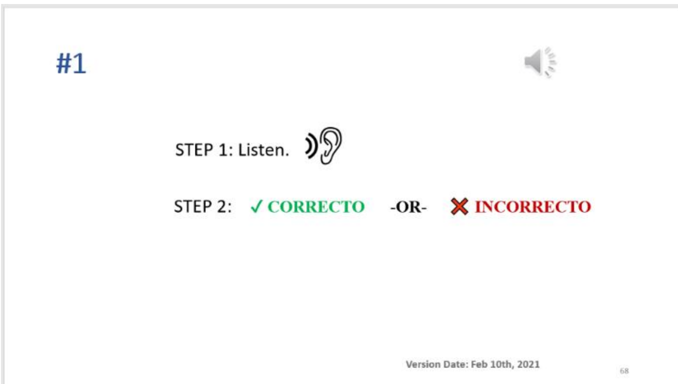


Task 3: Speeded Auditory Grammaticality Judgement (speeded AGJT)

This task consisted of 32 tokens total (16 targets; 16 distractors) presented aurally and consisting of simple statements in Spanish that featured both gender agreement and agreement violations (8 grammatical; 8 ungrammatical). Just like in Task 2, both the grammatical and ungrammatical tokens featured an even distribution between both error/focus types (8 Det-N and 8 N-Adj.) and avoided gendered cues that would reveal noun gender assignment by including only adjectives that are invariable for gender in tokens with Det-N focus and either the invariable determiner *cada* ("each") or an invariable possessive determiner in tokens with N-Adj error/focus. The target tokens were evenly distributed by gender class (8 masculine; 8 feminine) and noun morphology (8 overtly marked/canonical; 8 non-overtly marked/noncanonical) and were presented in randomized order. The distractor tokens consisted of verb phrases containing subject-verb agreement and violations. Like Task 2, just one statement was presented per PowerPoint slide with a small speaker icon which the researcher pressed to play the audio recording. Participants were instructed to listen to the statements all recorded by the same male native Spanish speaker and then indicate orally if the statement was grammatically correct or incorrect as prompted in the PowerPoint slide (e.g., ungrammatical, N-Adj. error: *mi clase es pequeño*, "my-invar. class-fem. is small-masc.; grammatical, Det-N focus: *veo una casa grande*, "I see a-fem. big-invar. house-fem"). This task began with two practice sets that featured non-target stimuli (e.g., ungrammatical: *me gustan correr en el parque*, "I like-pl. to run in the park-sg."; e.g., grammatical*: ella quiere vivir en el centro*, "she-3rd p.sg. wants-3rd p.sg. to live in downtown"). Like Task 2, this task was also speeded in that participants were instructed to respond as quickly as possible with their first impression. See Example 2 below for the task instructions participants saw and an example prompt. A full list of the stimuli used in this task is presented in *Appendix D*.

Example 2. Task 3 (Speeded Auditory Grammaticality Judgment Task) instructions and sample prompt.

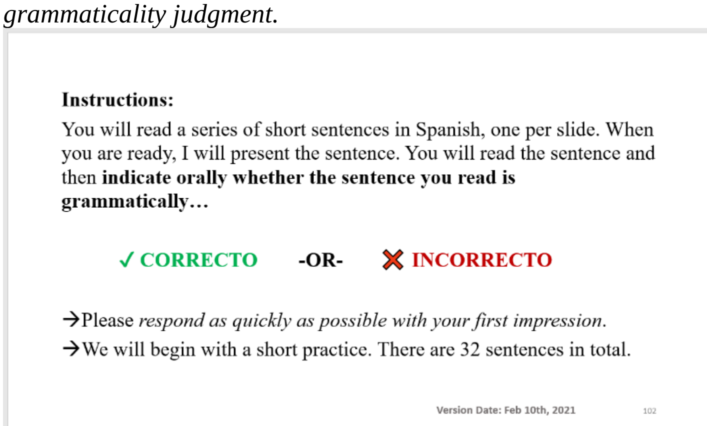


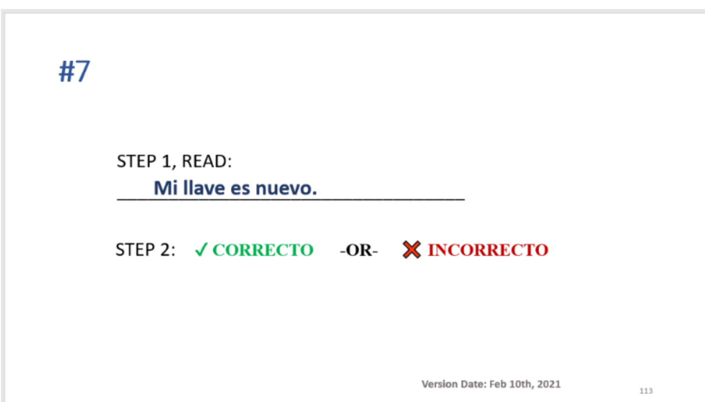


Task 4: Speeded Written Grammaticality Judgement (speeded WGJT)

This task consisted of 32 tokens total (16 targets; 16 distractors), presented in randomized order, with the same type of simple grammatical and ungrammatical statements in Spanish as Task 3, contrasting both focus and error types (8 grammatical: 4 Det-N focus; 4 N-Adj. focus; 8 ungrammatical: 4 Det-N error; 4 N-Adj. error), but this time presented in the written modality. The target tokens were also evenly distributed by gender class (8 masculine; 8 feminine) and noun morphology (8 overtly marked/canonical; 8 non-overtly marked/noncanonical). Just like in Task 3, the distractor tokens consisted of verb phrases containing subject-verb agreement and violations. Like the previous tasks, just one statement was presented per PowerPoint slide, but instead of a speaker icon, the written sentence in Spanish appeared on each slide. In order to ensure that participants were moving through the prompts as quickly as possible and reading each sentence only one time, they were instructed to read each statement *out loud* as quickly as possible and to indicate orally if the statement was grammatically ‘correct’ or ‘incorrect’, as prompted in the PowerPoint slide (e.g., ungrammatical, N-Adj. error: *tu ventana está abierto*, “your-invar. window-fem. is open-masc.”; grammatical, Det-N focus: *veo un escritorio verde*, “I see a-masc. green-invar. desk-masc.”). This task also began with two practice sets that featured non-target stimuli (e.g., ungrammatical: *nosotros comen mucho arroz*, “we-2nd p.pl. eat-3rd p.pl. a lot of rice”; grammatical: *él está en la biblioteca*, “he-3rd p.sg. is-3rd p.sg. in the library”). Like the previous tasks, this task was also speeded such that participants were instructed to respond as quickly as possible with their first impression. See Example 3 below for the task instructions participants saw and an example prompt. A full list of the stimuli used in this task is provided in *Appendix E*.

Example 3. Task 4 (Speeded Written Grammaticality Judgment Task) instructions and sample prompt.
**Note: participants were verbally instructed to read each prompt out loud before responding with their grammaticality judgment.*

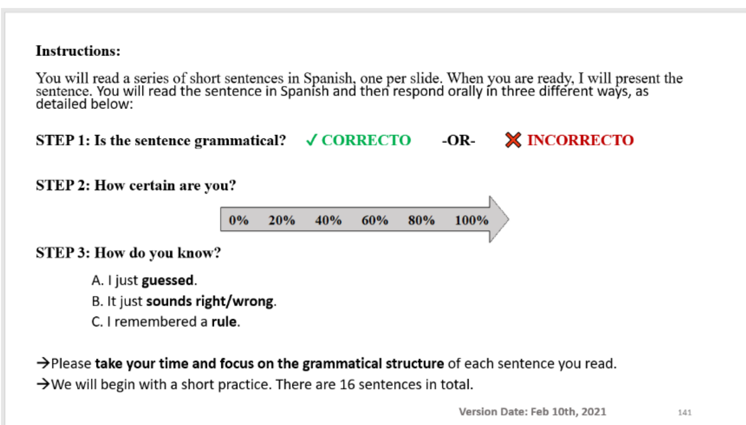


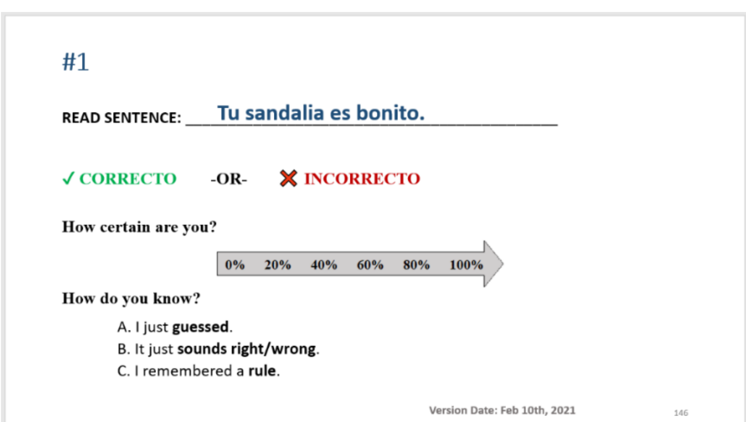


Task 5: Self-paced Written GJT

This task consisted of 16 tokens total (all targets; no distractors), half grammatical and half ungrammatical (8 grammatical: 4 Det.-N focus; 4 N-Adj. focus; 8 ungrammatical: 4 Det.-N errors; 4 N-Adj. errors) that were already presented in Task 3 and Task 4 (8 tokens from AGJT; 8 tokens from WGJT), and which therefore have already been distributed equally across both gender classes and morphological marking types, and were presented in randomized order. Each sentence was provided on an individual slide in the written modality. Participants were instructed to read each sentence provided and respond orally in three ways, as indicated in the PowerPoint slide for each token: 1. indicate whether each sentence is grammatical or ungrammatical; 2. indicate the degree of certainty of their judgement (as proposed by Sorace, 1996) along a percentage scale in six increments of twenty percent; and 3. self-report the source attribute for their response by selecting one of three options provided (a.) I'm not sure; I just guessed, b.) It just sounds right/wrong, or c.) I remembered a rule.). This task was self-paced, so participants were instructed to take their time. The idea behind the self-paced nature of this task was to condition participants to access their explicit knowledge about language structure while focusing on form. See Example 4 below for the task instructions participants saw and an example prompt. A full list of the stimuli used in this task is provided in *Appendix F*.

Example 4. Task 5 (Self-paced Written Grammaticality Judgment Task) instructions and sample prompt.

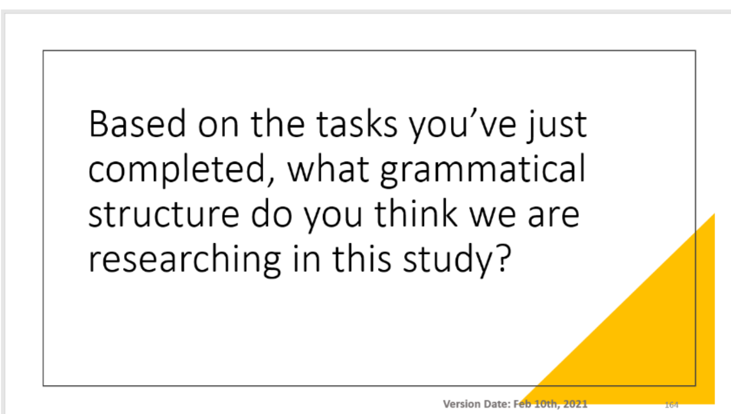


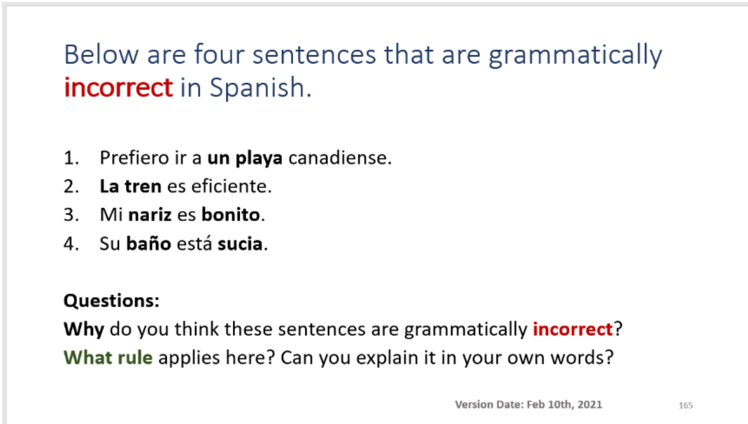


Task 6: Metalinguistic Awareness Exit Survey

At the end of the Zoom session, participants were asked by the researcher, in English for nonnative learners and in Spanish for native Spanish speakers, to give their best guess in their own words of what they thought was being researched in the study based on the tasks they had just performed. If participants were able to accurately identify the study's objective (e.g., gender, agreement, how to combine articles, nouns, and adjectives in Spanish, etc.), they were then be prompted to explain the grammatical rule to the best of their ability in their own words (i.e., no technical terminology needed). However, if they were not able to articulate the perceived research objective and/or if they were not able to generate a grammatical rule, they were provided with four sample ungrammatical sentences in Spanish (2 agreement violations with masculine nouns; 2 agreement violations with feminine nouns; one per error type: Det-N error; N-Adj. error) taken directly from the experimental tasks. The researcher then stated that the shown sentences are ungrammatical and asked the participant to explain why they thought the sentences are ungrammatical and, if possible, to provide some sort of rule. See Example 5 below for the two prompts that participants saw during the exit survey. A full list of the stimuli used in this task is provided in *Appendix G*.

Example 5. Task 6 (Metalinguistic Awareness Exit Survey) prompt.





3.4 Analysis: variables

All values for independent and dependent variables were entered into an Excel spreadsheet. Inferential analysis was carried out in *Jamovi* (version 2.3.18), which is an open-access and open-source statistical package that utilizes R script in a user-friendly click-based platform. Table 7 outlines the dependent and independent variables that were analyzed during performance with grammatical gender in Spanish.

Table 7. Outline of study variables of analysis and respective categories.

Dependent/ response variables:	▪ ***accuracy*** rate
	▪ inter- (SD per group) and intra- speaker ***variation*** (variance per individual)
Independent/ explanatory variables:	
Task variables (task demands)	▪ ***stimuli modality***: aural vs. written (audio only prompts vs. written only prompts) ▪ ***time constraint:*** speeded (time constrained; favors implicit knowledge) vs. self-paced (not time-constrained; conditions use of explicit knowledge)
Linguistic variables	▪ ***noun gender class*** (masculine/feminine) ▪ ***noun morphology*** (overt/non-overt morphology for gender) ▪ ***domain of agreement*** (Det.-N vs. N-Adj. agreement) ▪ ***target noun frequency*** (from Corpus de Referencia del Español Actual, CREA)
Individual variables	▪ ***speaker status***/context of acquisition: native/naturalistic vs. nonnative/instructed ▪ ***motivation*** type and level as characterized within Dörnyei's (2005) L2MSS framework ▪ ***attitudes*** about language learning and target language community ▪ ***metacognitive awareness*** level ▪ ***metalinguistic awareness*** level of target structure ▪ ***proficiency*** (tested & self-report)

	▪ ***multilingualism***/multilingual experience/exposure (both number and nature of languages known – i.e., presence of grammatical gender in prior linguistic repertoire)

The raw data for each task was entered into an Excel spreadsheet such that each prompt constituted a row, and each participant constituted a column. All participant responses on the Spanish proficiency test and the experimental tasks (Tasks 2-5) were coded in a binary fashion such that a correct response received a score of '1' and an incorrect response received a score of '0'. Task 6, the Metalinguistic Awareness Exit Survey, was also coded in a binary fashion in which participants were either overtly "aware" (1) or "not aware" (0) of the grammatical structure being tested, and if they were not overtly aware, if that awareness could be conditioned by the researcher during the exit survey (yes: '1' or no: '0'), producing a binary score for each participant on both overt and conditioned metalinguistic awareness of grammatical gender in Spanish.

After all raw data was entered and coded in Excel, summary cells were constructed with the average scores on each independent variable analyzed, including scores on: target tokens, distractor tokens, grammatical tokens, ungrammatical tokens, masculine tokens, feminine tokens, overtly-marked tokens, non-overtly-marked tokens, determiner-focused tokens, adjective-focused tokens, in addition to each participant's average score on ungrammatical feminine noun targets and ungrammatical feminine non-overt noun targets, as the two token types that are most indicative of the real acquisition of grammatical gender since they are less susceptible to default responses (i.e., "grammatical", "masculine"). In addition, average scores on high frequency and low frequency target nouns were calculated, based on the frequency scores reported in the *Corpus de Referencia del Español Actual* (CREA)[2].

The summary data was then transposed into a separate Excel spreadsheet in which each column constituted a dependent response variable of analysis, including accuracy scores on each task and on each task type (auditory vs. written; speeded vs. self-paced) and variation in the form of standard deviation (SD), and each row represented a participant. Experimental task totals were

[2] For relative noun frequencies in the Spanish language as tabulated and reported by the *Real Academia Española* in the Corpus de Referencia del Español Actual (CREA), the following webpage was consulted: https://corpus.rae.es/lfrecuencias.html

calculated for each participant and totals on the other independent individual variables, including speaker status (native vs. nonnative), L1, Spanish linguistic history score, Spanish use score, multilingualism (yes/no, number of additional languages), Spanish attitude score, as well as self-reported and tested Spanish proficiency scores were calculated for each participant and, in addition, metacognitive awareness score and motivation score were calculated for the adult learner participants only. Recall that the scoring procedure for each individual variable was previously described in the *Materials and Tasks section* under the Language Learner Profile Questionnaire.

Both descriptive and inferential data analysis were conducted in Jamovi (version: 2.3.18). Paired samples t-tests were used as an initial measure of significance for the effect on accuracy scores of each binary linguistic variable analyzed as well as for task effects comparing average accuracy scores on speeded vs. self-paced tasks and on auditory vs. written stimuli tasks. Correlation matrices were also constructed producing a Pearson's r correlation coefficient to examine the relationship between each individual variable (proficiency, metacognitive awareness, motivational orientation, metalinguistic awareness, Spanish use, Spanish attitudes, multilingualism, and prior linguistic repertoire) and average accuracy scores on each task type and overall. Finally, multiple linear regression modeling was used to determine which independent variables most significantly impacted accuracy in performance by constructing separate models for the combined impact of the task variables (time constraint, stimuli modality), linguistic variables (target noun gender class, morphology, domain of agreement, and relative target noun frequency), and individual variables (motivation, attitudes, metacognitive and metalinguistic awareness, Spanish proficiency, Spanish use, and multilingualism factors) for native Spanish speakers and adult learners separately. This allowed us to examine both the predictive power of each regression model (expressed as adjusted R^2 and F-value) as well as the relative contribution of each model coefficient (F-value) and its corresponding significance to the model (p-value).

The following two chapters present both the descriptive and inferential results for each research question posed. Chapter 4 focuses on the results of the linguistic and task effects (questions 1-2) and Chapter 5 focuses on the results of the individual factors and the multilingual effect (questions 3-4).

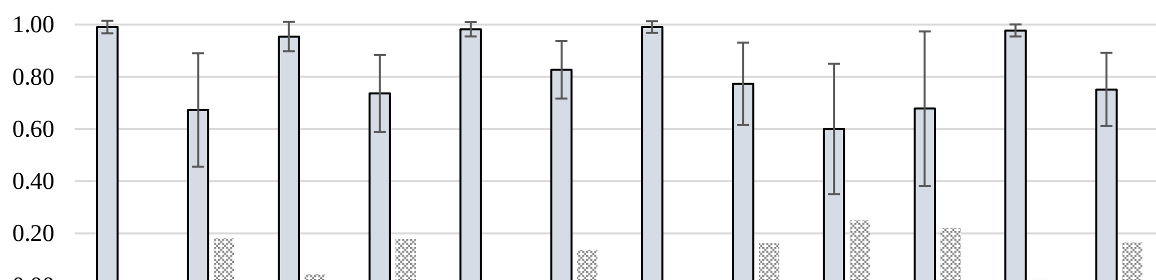
Chapter 4: Results: Linguistic and task effects

Data collected from the Language Learner Profile Questionnaire and the Spanish proficiency test (Task 1) are provided in the Methodology section. The results from each of the four experimental tasks (Task 2-5) and Task 6, the metalinguistic awareness exit survey, are presented below, according to the first two research questions posed concerning the relative effect of each linguistic variable as well as how task factors impact performance with grammatical gender in Spanish.

4.1 Performance with gender agreement

The first research question examined how native Spanish speakers and adult/late Spanish learners perform with grammatical gender agreement and how the linguistic variables of noun gender class (masculine vs. feminine), morphological marking (overtly marked vs. non-overtly marked), domain of agreement (Det-N vs. N-Adj), and target noun frequency (high vs. low relative frequency) may modulate participants' performance in terms of both accuracy and variation. Native Spanish speaker (n = 25) and adult Spanish learner (n = 90) participants' accuracy scores and variation in performance (both as a group measured as standard deviation and individual variation measured as Variance) were compared on each of the experimental tasks (Tasks 2-5) and on the exit survey (Task 6). In what follows, the descriptive results will be presented in detail for each linguistic factor analyzed and then the results of inferential analysis using *t*-tests and multiple linear regression modeling will be presented to substantiate the descriptive observations.

As Figure 13 shows, native speakers, as expected, performed at or near ceiling on most experimental tasks, but notably did not perform at ceiling on the metalinguistic awareness exit survey (Task 6), in which the average native speaker score was only 60%; although native speakers demonstrated very high accuracy with grammatical gender, they were not as explicitly aware of the fact that grammatical gender was being tested nor as able to explain what grammatical gender is and the rules governing it. Learners, on the other hand, performed considerably below their native-speaker peers (Tasks 2-5), despite being, on average, much more explicitly aware of the target structure and more able to explain its use with ease (as demonstrated on Task 6). The experimental task that resulted in the greatest difference in accuracy scores between the two main participant groups was Task 2, the speeded oral imitation task, featuring auditory only stimuli, in which the average native speaker accuracy rate was 31.7 points higher than the average of adult learner accuracy rate. The second most difficult task for adult Spanish learners was Task 3, the speeded auditory grammaticality judgement task, producing a 21.8% difference in accuracy between learners and native speakers. Learner participants exhibited their highest accuracy rate on Task 4, the speeded written grammaticality judgement task (82.7%), while native speakers performed best on the speeded oral imitation task (Task 2). In summary, the average native speaker accuracy rate on all experimental tasks combined was 97.8% while the average learner accuracy rate was 75.2%, as demonstrated in Figure 13. Regarding individual variation in performance (measured as Variance), native speakers consistently demonstrated very low intra-speaker variance in their scores, with minor fluctuations on Task 3 and Task 6, while learners demonstrated consistently higher variation across all tasks, both as a group (observable in the standard deviation bars) and individually (demonstrated in the intra-learner/speaker Variance bars in grey), as shown in Figure 13.



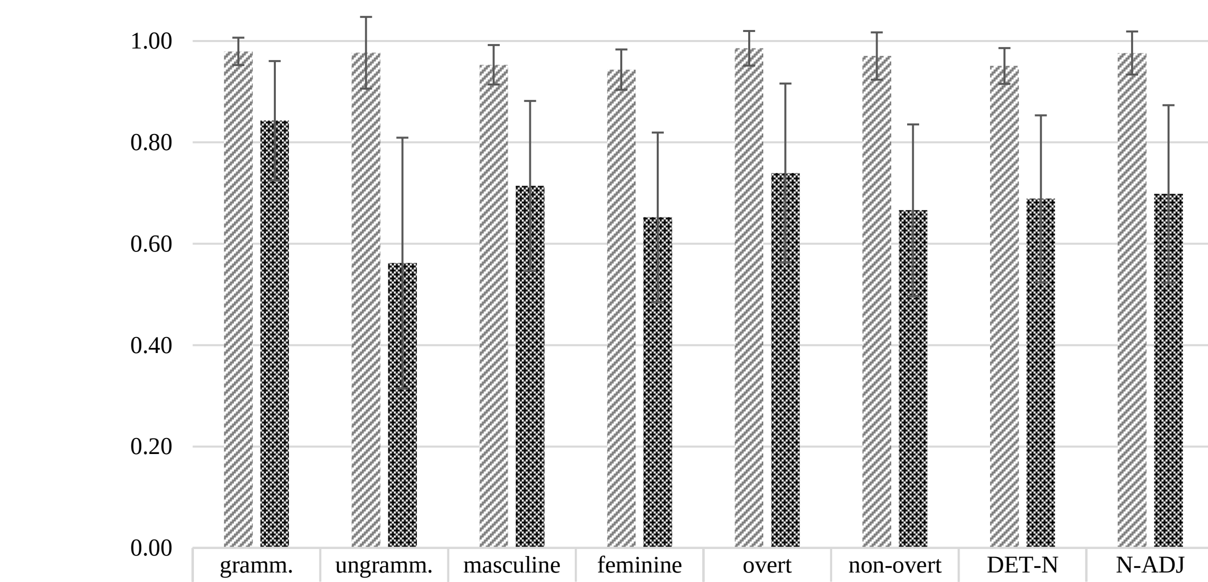
	TASK 2		TASK 3		TASK 4		TASK 5		TASK 6		AVERAGE	
	Native	Learner	Native	Learner	Native	Learner	Native	Learner	Native	Learner	Native	Learner
Score	0.990	0.673	0.954	0.736	0.982	0.827	0.990	0.773	0.600	0.678	0.978	0.752
Variance	0.010	0.181	0.043	0.179	0.018	0.136	0.010	0.162	0.250	0.221	0.022	0.165

Score **Variance**

Figure 13. Accuracy scores and intra-learner/speaker variance compared per group per task: Task 2 (Speeded Oral Imitation Task – auditory), Task 3 (Speeded Auditory Grammaticality Judgement Task), Task 4 (Speeded Written Grammaticality Judgement Task), Task 5 (Self-paced Written Grammaticality Judgement Task), and Task 6 (Metalinguistic Awareness Exit Survey). Standard deviation (SD) bars are provided.

4.1.1 Effect of linguistic variables

Regarding the relative effect of each of the linguistic variables investigated, Figure 14 demonstrates the mean and standard deviation (represented as error bars) per participant group on each binary variable analyzed, including token grammaticality (grammatical/ungrammatical), noun gender class (masculine/feminine), noun morphology (overt/non-overt), and domain of agreement (Det-N/N-Adj). For the purpose of this preliminary analysis, adult learners have been grouped together and compared to their native speaker counterparts in order to examine more broadly how the linguistic variables inherent to grammatical gender in Spanish may differentially affect both native speakers and late/adult learners. The effect of each linguistic variable per learner proficiency level will be presented later on in this section (see Figure 16). Overall, while native speaker accuracy scores remained relatively consistent across variables, clear distinctions are observed in accuracy scores in adult learners who exhibited enhanced accuracy with grammatical, masculine, overtly marked tokens. Data summary tables, organized per task (Task 2-Task 6), including accuracy scores, intra-learner/speaker variance, group SD, and scores per linguistic variable, can be found in the *Appendices* (*Appendices* H-L).



	Token Grammaticality		Noun Gender Class		Noun Morphology Type		Domain of Agreement	
	gramm.	ungramm.	masculine	feminine	overt	non-overt	DET-N	N-ADJ
Native Spanish	0.98	0.98	0.95	0.94	0.99	0.97	0.95	0.98
Spanish Learner	0.84	0.56	0.71	0.65	0.74	0.67	0.69	0.70

Figure 14. Effect of linguistic variables compared per group. Standard deviation (SD) bars are provided.

In all, although the average total task score (on all tasks combined) was 97.8% for native speakers compared to 75.2% for learners, as shown in Table 8, the linguistic variables analyzed produced a much larger effect on accuracy scores in the learner group, whereas native speakers were largely unaffected by these variables. The two main participant groups were most distinct from one another on their accuracy with ungrammatical, feminine, non-overtly marked tokens with an average learner group accuracy score of 41.2%, whereas their native speaker counterparts scored a 96.1% on these same tokens when averaged across tasks.

Table 8. Experimental task totals descriptive data summary for all participants.

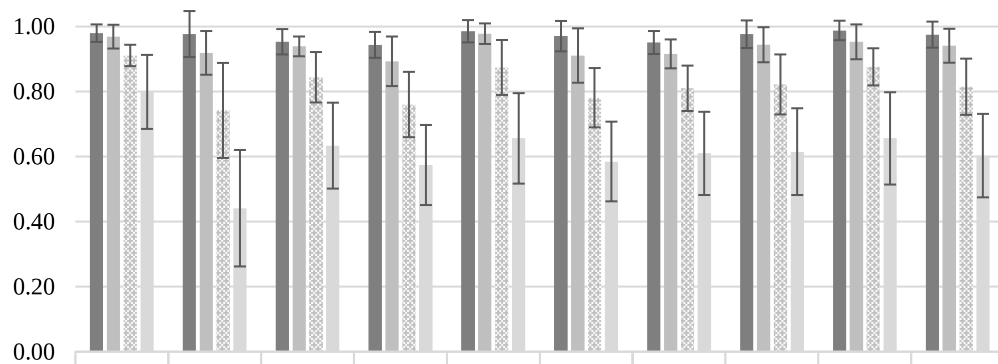
	SPEAKER STATUS	N	Mean	SD
AVG TOTAL TASK SCORE	Native Spanish	25	0.978	0.023
	Spanish Learner	90	0.752	0.140
AVG GRAMM SCORE	Native Spanish	25	0.980	0.027
	Spanish Learner	90	0.843	0.118
AVG UNGRAMM SCORE	Native Spanish	25	0.977	0.071
	Spanish Learner	90	0.562	0.247
AVG MASC. SCORE	Native Spanish	25	0.953	0.039
	Spanish Learner	90	0.714	0.168
AVG FEM. SCORE	Native Spanish	25	0.944	0.040
	Spanish Learner	90	0.652	0.167
AVG OVERT SCORE	Native Spanish	25	0.986	0.034
	Spanish Learner	90	0.740	0.176
AVG NON-OVERT SCORE	Native Spanish	25	0.970	0.047
	Spanish Learner	90	0.666	0.169
AVG DET-N SCORE	Native Spanish	25	0.951	0.035
	Spanish Learner	90	0.689	0.165
AVG N-ADJ SCORE	Native Spanish	25	0.976	0.043
	Spanish Learner	90	0.698	0.175
Avg. Ungramm. Fem. Score	Native Spanish	25	0.975	0.081
	Spanish Learner	90	0.509	0.259
Avg. Ungramm. Fem. Non-overt Score	Native Spanish	25	0.961	0.112
	Spanish Learner	90	0.412	0.288

Furthermore, the effect of target noun frequency was analyzed and compared for each participant group. Table 9 presents the descriptive results of the target noun frequency analysis with the mean accuracy scores on all experimental tasks combined per participant group on high frequency nouns, low frequency nouns, the frequency differential score (i.e., the average difference in accuracy between high frequency and low frequency nouns), as well as the correlation coefficient (Pearson's *r*) between relative noun frequency and target noun accuracy scores. As we can observe depicted graphically in Figure 15, native speakers were only slightly affected by noun frequency (1.3% difference between high and low frequency accuracy scores), whereas learners were much more affected, demonstrating an average difference in accuracy scores of 4.7% between high and low frequency target nouns, with a correspondingly higher correlation coefficient of .06 for all learners combined versus only .02 for native speakers. Nonetheless, the correlations between accuracy and noun frequency are notably weak in both groups and the effect of target noun frequency appears to be modulated by Spanish proficiency level in nonnative learners, as demonstrated in Figure 15.

Table 9. Target noun frequency analysis for all participants on all tasks combined.

	SPEAKER STATUS	N	MEAN	SD
HIGH FREQ. SCORE	Native Spanish	25	0.988	0.030
	Spanish Learner	90	0.736	0.172
LOW FREQ. SCORE	Native Spanish	25	0.975	0.040
	Spanish Learner	90	0.689	0.174
FREQ. DIFF. SCORE	Native Spanish	25	0.012	0.027
	Spanish Learner	90	0.046	0.087
FREQ-SCORE CORRELATION (*r*)	Native Spanish	25	0.019	0.069
	Spanish Learner	90	0.062	0.079

In order to examine to what extent proficiency in Spanish may influence the relative effect on performance of the linguistic variables analyzed (token grammaticality, noun gender class, noun morphology type, domain of agreement, noun frequency), accuracy scores were compared across learner proficiency groups and the native speaker group. As shown in Figure 15, advanced proficiency learners patterned very closely with their native speaker counterparts on all linguistic measures analyzed. Although the advanced proficiency learner group demonstrated *qualitatively* the same pattern in their performance as the lower proficiency groups (i.e., intermediate and beginner), with enhanced accuracy on grammatical, masculine, overtly marked tokens, *quantitatively* their scores were more similar to the native speaker group than to the intermediate learner group, thereby demonstrating that as proficiency increases, the effect that linguistic variables have on learner performance diminishes and advanced adult learners can very closely approximate native speaker norms of performance.



	gramm.	ungram.	masc.	fem.	overt	non-overt	DET-N	N-ADJ	high	low
	Token Grammaticality		Noun Gender Class		Noun Morphology Type		Domain of Agreement		Noun Frequency	
NATIVE (n=25)	0.98	0.98	0.95	0.94	0.99	0.97	0.95	0.98	0.99	0.98
ADVANCED (n=14)	0.97	0.92	0.94	0.89	0.98	0.91	0.92	0.94	0.95	0.94
INTERMEDIATE (n=14)	0.91	0.74	0.84	0.76	0.87	0.78	0.81	0.82	0.88	0.82
BEGINNER (n=62)	0.80	0.44	0.63	0.57	0.66	0.59	0.61	0.62	0.66	0.60

Figure 15. Effect of linguistic variables (token grammaticality, noun gender class, noun morphology type, domain of agreement, and noun frequency) compared per Spanish proficiency group and native speakers. Standard deviation (SD) bars are provided.

Paired samples *t*-tests, with a confidence interval of 95%, were performed to examine the significance of the relative effect on mean accuracy scores of each of the binary linguistic variables analyzed as well as noun frequency. Table 10 presents the results for the adult learners and Table 11 presents the results for the native speakers. With regards to the learner group, significant differences in mean accuracy were detected for grammatical/ungrammatical tokens, masculine/feminine tokens, overtly marked/non-overtly marked tokens, as well as high frequency/low frequency tokens, all resulting in medium to large effect sizes in the following decreasing order: grammatical (M = .843, SD = .117) as compared to ungrammatical tokens (M =

.562, SD = .247), $t(89) = 13.66$, $d = 1.440$, $p < .001$; overtly-marked nouns (M = .740, SD = .176) as compared to non-overtly marked nouns (M = .666, SD = .169), $t(89) = 7.68$, $d = .809$, $p < .001$; masculine nouns (M = .714, SD = .168) as compared to feminine nouns (M = .652, SD = .167), $t(89) = 6.24$, $d = .657$, $p < .001$; and finally high-frequency nouns (M = .736, SD = .172) as compared to low-frequency nouns (M = .689, SD =.174), $t(89) = 5.22$, $d = .551$, $p < .001$. Interestingly, native Spanish speakers also showed enhanced accuracy on certain token types, although to a lesser extent. As shown in Table 10, native speakers demonstrated significantly higher mean accuracy rates with overtly marked nouns (M = .986, SD = .034) than non-overtly marked nouns (M = .970, SD = .047), $t(24) = 2.96$, $d = .592$, $p = .007$, and higher frequency nouns (M = .988, SD = .030) than lower frequency nouns (M = .975, SD = .040), $t(24) = 2.31$, d = .461, $p = .030$, although there was no significant difference detected in their accuracy for grammatical vs. ungrammatical tokens nor for masculine vs. feminine nouns. Contrary to their adult learner peers, native Spanish speakers did show a significant effect of domain of agreement in which they were more accurate with N-Adj agreement (M = .976, SD = .043) than with Det-N agreement (M = .951, SD = .035), $t(24) = -4.84$, $d = -.967$, $p < .001$. In summary, the average accuracy scores of both adult learners and native speakers were affected by the linguistic variables analyzed in which both groups demonstrated a significant effect of target noun morphology type (i.e., overtly vs. non-overtly marked) and target noun frequency, although the effect size (expressed as Cohen's d) was larger in the learner group. In addition, adult learners also demonstrated a significant effect of token grammaticality (i.e., grammatical vs. ungrammatical) as well as noun gender class (i.e., masculine vs. feminine), while the performance of native speakers did not appear to be significantly affected by these factors.

Table 10. Results of a paired samples *t*-test examining the relative effect of each linguistic variable and target noun frequency in **learners** only. Significant (95% CI) differences are highlighted in grey.

		t	df	*p*	Mean difference	95% Confidence Interval: Lower	95% Confidence Interval: Upper	Cohen's d Effect Size
AVG GRAMMATICAL SCORE	**AVG UNGRAMMATICAL SCORE**	13.66	89.0	< .001	0.28089	0.2400	0.32175	1.440
AVG MASC SCORE	**AVG FEM SCORE**	6.24	89.0	< .001	0.06189	0.0422	0.08161	0.657
AVG OVERT SCORE	**AVG NON-OVERT SCORE**	7.68	89.0	< .001	0.07344	0.0544	0.09245	0.809
AVG DET-N SCORE	**AVG N-ADJ SCORE**	-1.02	89.0	0.311	-0.00956	-0.0282	0.00910	-0.107
NOUN FREQ. SCORE: HIGH	**NOUN FREQ. SCORE: LOW**	5.22	89.0	< .001	0.04778	0.0296	0.06595	0.551

Table 11. Results of a paired samples *t*-test examining the relative effect of each linguistic variable and target noun frequency in **native Spanish speakers** only. Significant (95% CI) differences are highlighted in grey.

		t	df	*p*	Mean difference	95% Confidence Interval: Lower	95% Confidence Interval: Upper	Cohen's d Effect Size
AVG GRAMMATICAL SCORE	**AVG UNGRAMMATICAL SCORE**	0.185	24.0	0.855	0.00280	-0.02842	0.0340	0.0370
AVG MASC SCORE	**AVG FEM SCORE**	1.788	24.0	0.086	0.00960	-0.00148	0.0207	0.3576
AVG OVERT SCORE	**AVG NON-OVERT SCORE**	2.960	24.0	0.007	0.01520	0.00460	0.0258	0.5920
AVG DET-N SCORE	**AVG N-ADJ SCORE**	-4.836	24.0	< .001	-0.02560	-0.03653	-0.0147	-0.9671
NOUN FREQ. SCORE: HIGH	**NOUN FREQ. SCORE: LOW**	2.305	24.0	0.030	0.01280	0.00134	0.0243	0.4610

4.1.2 Effect of linguistic variables: predictive modeling

In order to investigate how the linguistic factors analyzed together impact performance with grammatical gender in both native Spanish speakers and adult Spanish learners, multiple linear regression models were constructed at a confidence interval of 95%. The final model demonstrating the impact of linguistic factors in adult learners is provided in Table 12, while Table 13 shows the model for native Spanish speakers.

A significant multiple linear regression model was found that accounts for approximately 97% of the variation in average task scores in adult Spanish learners (adjusted $R^2 = 0.966$, $F = 637$, $p <$

.001), including the linguistic factors of token grammaticality (*ungrammatical*), noun gender class (*feminine*), morphological class (*non-overt*), and frequency group (*low frequency*). When the model coefficients are considered individually, average ungrammatical score (F = 19.94, p < .001), average feminine noun score (F = 12.97, p < .001), and low-frequency noun score (F = 109.65, p < .001) are all individually significant factors; however, average non-overt score (F = 0.334, p = .565) is a non-significant factor when considered separately, according to the output of an Omnibus ANOVA test, as shown in Table 12. Therefore, in late Spanish learners, relative noun frequency is most predictive of performance with grammatical gender, followed by token grammaticality and noun gender class, whereas noun morphology does not appear to be significantly predictive of performance when all linguistic factors are considered together.

Table 12. Output of a multiple linear regression model (CI 95%) examining the effects of the linguistic variables on average task scores, including token grammaticality (ungrammatical), noun gender class (feminine), morphological class (non-overt), and frequency group (low frequency) in **learners** of Spanish.

Model Fit Measures

				Overall Model Test			
Model	**R**	**R^2**	**Adjusted R^2**	***F***	**df 1**	**df 2**	***p***
1	0.984	0.968	0.966	637	4	85	< .001

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
AVG UNGRAMMATICAL SCORE	0.01322	1	0.01322	19.936	< .001
AVG FEM SCORE	0.00860	1	0.00860	12.968	< .001
AVG NON-OVERT SCORE	2.21e-4	1	2.21e-4	0.334	0.565
NOUN FREQ. SCORE: LOW	0.07273	1	0.07273	109.645	< .001
Residuals	0.05638	85	6.63e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

Predictor	Estimate	SE	95% Confidence Interval Lower	95% Confidence Interval Upper	*t*	*p*
Intercept	0.1343	0.0196	0.0953	0.173	6.839	< .001
AVG UNGRAMMATICAL SCORE	-0.1858	0.0416	-0.2685	-0.103	-4.465	< .001
AVG FEM SCORE	0.1969	0.0547	0.0882	0.306	3.601	< .001
AVG NON-OVERT SCORE	0.0345	0.0597	-0.0842	0.153	0.578	0.565
NOUN FREQ. SCORE: LOW	0.8281	0.0791	0.6709	0.985	10.471	< .001

A significant multiple linear regression model was also found that accounts for approximately 82% of the variation in average task scores in native Spanish speakers (adjusted $R^2 = 0.822$, $F = 23.2$, $p < .001$), including the linguistic factors of token grammaticality (*ungrammatical*), noun gender class (*feminine*), morphological class (*non-overt*), domain of agreement (*Det-N*), and frequency group (*low frequency*). However, none of the individual model coefficients reached significance, according to the output of an Omnibus ANOVA test, as shown in Table 13. Therefore, although when considered all together, the linguistic factors analyzed appear to be largely predictive of performance, none of the linguistic factors are *individually* predictive of performance with grammatical gender in native Spanish speakers.

Table 13. Output of a multiple linear regression model (CI 95%) examining the effects of the linguistic variables on average task scores, including token grammaticality (ungrammatical), noun gender class (feminine), morphological class (non-overt), domain of agreement (Det-N), and frequency group (low frequency) in **native speakers** of Spanish.

Model Fit Measures

				Overall Model Test			
Model	**R**	**R^2**	**Adjusted R^2**	***F***	**df 1**	**df 2**	***p***
1	0.927	0.859	0.822	23.2	5	19	< .001

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
AVG UNGRAMMATICAL SCORE	8.20e-7	1	8.20e-7	0.00863	0.927
AVG FEM SCORE	2.14e-4	1	2.14e-4	2.25202	0.150
AVG NON-OVERT SCORE	9.72e-5	1	9.72e-5	1.02172	0.325
AVG DET-N SCORE	2.57e-5	1	2.57e-5	0.27061	0.609
NOUN FREQ. SCORE: LOW	2.66e-4	1	2.66e-4	2.79804	0.111
Residuals	0.00181	19	9.51e-5		

Note. Type 3 sum of squares

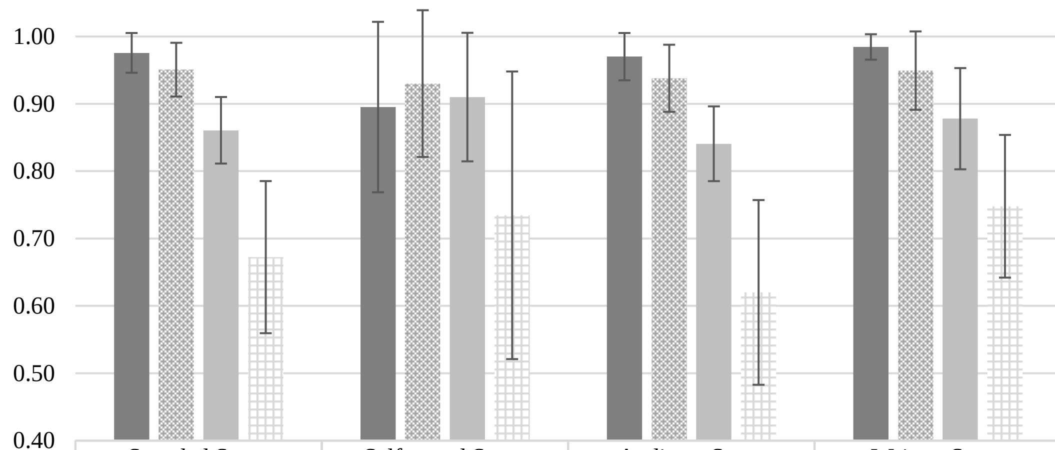
Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept	0.40872	0.0781	0.245	0.572	5.2304	< .001
AVG UNGRAMMATICAL SCORE	-0.00900	0.0969	-0.212	0.194	-0.0929	0.927
AVG FEM SCORE	0.26461	0.1763	-0.104	0.634	1.5007	0.150
AVG NON-OVERT SCORE	-0.20876	0.2065	-0.641	0.224	-1.0108	0.325
AVG DET-N SCORE	0.07459	0.1434	-0.226	0.375	0.5202	0.609
NOUN FREQ. SCORE: LOW	0.47134	0.2818	-0.118	1.061	1.6727	0.111

4.2 Task effects

The second research question examined how the task demands of a time constraint (speeded vs. self-paced) and stimuli presentation modality (auditory vs. written) impact accuracy and variation in the observed performance of both native Spanish speakers and adult learners of Spanish. Native Spanish speakers' (n=25) and adult learners' (n=90) accuracy scores were compared per task type in order to examine any task effects. The relative effect of a time constraint, that is, speeded as compared to self-paced tasks, as well as the relative effect of stimuli modality, that is, auditory as compared to written, were compared. In what follows, the descriptive results of the task effects analysis will be presented in detail and then the results of inferential analysis using *t*-tests and multiple linear regression modeling will be presented to substantiate the descriptive observations.

Figure 16 presents the descriptive results of this analysis. Learners were clearly affected by task type, although this effect varied per learner proficiency level. Intermediate (n=14) and beginner (n=62) proficiency learners performed notably better on self-paced and written modality tasks, yet their advanced proficiency learner peers (n=14) patterned very closely with the native speakers, both *qualitatively* by demonstrating greater accuracy on speeded tasks over self-paced tasks and no clear performance differences between auditory and written stimuli tasks, as well as *quantitatively* as their scores more closely approximated their native speaker counterparts than their other learner peers on speeded, auditory, and written average task scores. In contrast to the beginner and intermediate learners, the native speakers actually performed better on the speeded tasks and showed virtually no difference in their accuracy scores between written and auditory tasks. It is worth noting that both the advanced and intermediate learners outperformed the native speaker group on average scores with the self-paced tasks, as shown in Figure 16 below.



	Time Constraint		Input Modality	
	Speeded Score	Self-paced Score	Auditory Score	Written Score
Native (n=25)	0.98	0.90	0.97	0.98
Advanced (n=14)	0.95	0.93	0.94	0.95
Intermediate (n=14)	0.86	0.91	0.84	0.88
Beginner (n=62)	0.67	0.73	0.62	0.75

Native (n=25)   Advanced (n=14)   Intermediate (n=14)   Beginner (n=62)

Figure 16. Effect of task demands on accuracy scores compared per learner proficiency group and native speakers. Standard deviation (SD) bars are provided.

Regarding the possible effect of task demands on speaker variation in performance, native speakers' and adult learners' standard deviation can be compared across task types per group. As can be observed in the standard deviation bars in Figure 17, group variation steadily decreased with increasing proficiency level such that native speakers demonstrated the least variation per task type and the beginner learners showed the most group variation. On average self-paced task scores, all four groups (native speakers, advanced, intermediate, and beginner learners) demonstrated the most variation of all four task measures. Therefore, task type appears to have an impact on variation in performance such that both native speakers and learners of all proficiency levels are more variable in their responses on self-paced tasks yet the effect of task stimuli modality on variation in performance is less evident; only beginner learners tend to demonstrate more variation in their performance on auditory as compared to written stimuli tasks. Overall, the task effect of time constraint affects variation considerably more than the effect of stimuli modality in both native speaker and learner groups and beginner learner variation tends to be affected much more by task type such that self-paced and auditory tasks lead to the most variation in performance.

In order to investigate if the task demands effects detected in the present sample are significant, including the time constraint factor (speeded vs. self-paced) and the stimuli modality factor (auditory vs. written), paired samples t-tests (CI 95%) were performed. Table 14 presents the results for adult learners and Table 15 presents the results for the native speaker group. The task demands resulted in a significant mean difference in accuracy for both native speakers and adult learners. Learners demonstrated a significantly higher mean accuracy rate on the self-paced tasks (M = .792, SD = .204) compared to the speeded tasks (M = .745, SD = .147), producing a relatively small effect size ($t(89) = -2.60$, $d = -.274$, $p = .011$), and a significantly higher mean accuracy rate on the written tasks (M = .799, SD = .124) as compared to the auditory tasks (M = .704, SD = .174), producing a large effect size ($t(89) = -7.99$, $d = -.842$, $p < .001$). A similarly significant yet opposite trend is observed in the native speaker participant group in terms of the time constraint factor in which the native speakers demonstrated a significantly higher mean accuracy rate on the speeded tasks (M = .976, SD = .030) as compared to the self-paced tasks (M = .895, SD = .127), producing a medium effect size ($t(24) = 3.33$, $d = .666$, $p = .003$). With regards to stimuli modality, native speakers patterned similarly to the adult learners and demonstrated a significantly higher mean accuracy rate on the written tasks (M = .984, SD = .019) as compared to the auditory tasks (M = .970, SD = .035), although the effect size for this factor was relatively small ($t(24) = -2.35$, $d = -.469$, $p = .028$). In summary, task demands showed a significant effect on performance in both native speakers and adult learners. Native speakers showed, surprisingly, a larger task effect size (Cohen's d) for the time constraint factor than their learner peers in which they performed significantly better on speeded tasks while learners performed significantly better on self-paced tasks. Furthermore, both groups performed better on the written tasks, although the effect of stimuli presentation modality was most pronounced for the learners.

Table 14. Results of a paired samples *t*-test examining the effect of task demands including the time constraint factor (speeded vs. self-paced) and the input modality factor (auditory vs. written) for **learners** only. Significant (95% CI) differences are highlighted in grey.

						95% Confidence Interval		
		t	**df**	***p***	**Mean difference**	**Lower**	**Upper**	**Cohen's d Effect Size**
SPEEDED SCORE	**SELF-PACED SCORE**	-2.60	89.0	0.011	-0.0472	-0.083	-0.011	-0.274
AUDITORY SCORE	**WRITTEN SCORE**	-7.99	89.0	< .001	-0.0957	-0.120	-0.072	-0.842

Table 15. Results of a paired samples *t*-test examining the effect of task demands including the time constraint factor (speeded vs. self-paced) and the input modality factor (auditory vs. written) for **native Spanish speakers** only. Significant (95% CI) differences are highlighted in grey.

						95% Confidence Interval		
		t	**df**	***p***	**Mean difference**	**Lower**	**Upper**	**Cohen's d Effect Size**
SPEEDED SCORE	**SELF-PACED SCORE**	3.33	24.0	0.003	0.0804	0.031	0.1303	0.666
AUDITORY SCORE	**WRITTEN SCORE**	-2.35	24.0	0.028	-0.0144	-0.027	-0.0017	-0.469

4.2.1 Task effects: predictive modeling

In order to investigate how the task factors analyzed together impact performance with grammatical gender in both native Spanish speakers and adult Spanish learners, multiple linear regression models were constructed at a confidence interval of 95%. The final model demonstrating the impact of task factors in adult learners is provided in Table 16, while Table 17 shows the model for native Spanish speakers.

A significant multiple linear regression model was found that accounts for approximately 96% of the variation in average task scores in adult Spanish learners (adjusted $R^2 = 0.962$, $F = 1137$, $p < .001$), including the task factors of time constraint (*speeded*) and stimuli presentation modality (*auditory*). When the model coefficients are considered individually, average speeded score ($F = 108.3$, $p < .001$) and average auditory score ($F = 11.0$, $p = .001$) are individually significant factors, according to the output of an Omnibus ANOVA test, as shown in Table 16. Therefore, in adult Spanish learners, the time constraint task factor is most predictive of performance, followed by the stimuli modality task factor, and when both task factors are considered together, they are highly predictive of performance with grammatical gender in Spanish as a nonnative language.

Table 16. Output of a multiple linear regression model (CI 95%) examining the effects of the task factors on average task scores, including time constraint (speeded) and stimuli modality (auditory) in **learners** of Spanish.

Model Fit Measures

				Overall Model Test			
Model	**R**	**R^2**	**Adjusted R^2**	***F***	**df 1**	**df 2**	***p***
1	0.981	0.963	0.962	1137	2	87	< .001

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
SPEEDED SCORE	0.08022	1	0.08022	108.3	< .001
AUDITORY SCORE	0.00818	1	0.00818	11.0	0.001
Residuals	0.06442	87	7.40e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept	-0.00173	0.0234	-0.0483	0.0449	-0.0740	0.941
SPEEDED SCORE	1.35944	0.1306	1.0998	1.6190	10.4086	< .001
AUDITORY SCORE	-0.36872	0.1109	-0.5892	-0.1482	-3.3241	0.001

A significant multiple linear regression model was also found for native Spanish speakers that accounts for approximately 91% of the variation in average task scores (adjusted $R^2 = 0.906$, $F = 116$, $p < .001$), including the task factors of time constraint (*speeded*) and stimuli presentation modality (*auditory*). However, when the model coefficients are considered individually, only average speeded score ($F = 15.86$, $p < .001$) is significant, whereas average auditory score ($F = 0.02$, $p = .887$) is not an individually significant factor, according to the output of an Omnibus ANOVA test, as shown in Table 17. Therefore, in native Spanish speakers, the time constraint task factor is most predictive of performance; however, contrary to adult learners, native speakers perform significantly *better* under a time constraint and, furthermore, task stimuli modality is *not* independently predictive of performance with grammatical gender in Spanish as a native language, contrary to the trend observed in adult learners.

Table 17. Output of a multiple linear regression model (CI 95%) examining the effects of the task factors on average task scores, including time constraint (speeded) and stimuli modality (auditory) in **native speakers** of Spanish.

Model Fit Measures

				Overall Model Test			
Model	**R**	**R^2**	**Adjusted R^2**	***F***	**df 1**	**df 2**	***p***
1	0.956	0.914	0.906	116	2	22	< .001

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
SPEEDED SCORE	8.00e-4	1	8.00e-4	15.8593	< .001
AUDITORY SCORE	1.04e-6	1	1.04e-6	0.0207	0.887
Residuals	0.00111	22	5.05e-5		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept	0.2481	0.0533	0.138	0.359	4.659	< .001
SPEEDED SCORE	0.7259	0.1823	0.348	1.104	3.982	< .001
AUDITORY SCORE	0.0220	0.1529	-0.295	0.339	0.144	0.887

4.3 Summary of results: linguistic and task effects

In summary, with regards to the effect of the linguistic variables, although native speaker accuracy scores were largely unaffected by the linguistic factors, clear distinctions were observed in accuracy scores in adult learners who exhibited enhanced accuracy on frequent, grammatical, and masculine target tokens. Descriptive results indicated that native speakers were slightly affected by noun frequency while learners were much more affected. Proficiency level in adult learners was also found to interact with the relative effect of the linguistic variables analyzed as the advanced proficiency learner group demonstrated *qualitatively* the same pattern in their performance as the lower proficiency learner groups—showing greater accuracy with grammatical, masculine, and overtly marked tokens—yet *quantitatively* the advanced learner scores were more similar to the native speaker group than to the intermediate learner group as advanced learners approximated native speaker norms of performance. When considered all together through multiple linear regression modeling, relative noun frequency was found to be most predictive of learner performance with grammatical gender, followed by token grammaticality and noun gender class, whereas noun morphology did not appear to be independently predictive of performance in adult learners. In contrast, none of the linguistic factors were found to be individually predictive of performance with grammatical gender in Spanish as a native language.

Results also pointed to task effects, indicating that learners were clearly affected by task type and performed better on the self-paced tasks and on the written tasks, although relative task effects varied per learner proficiency level such that advanced learners patterned very closely—both quantitatively and qualitatively—with their native speaker counterparts. In contrast to the intermediate and beginner learner groups, both the native speakers and the advanced learners actually performed better on the speeded tasks and showed virtually no difference in their accuracy scores between written and auditory modality tasks. According to multiple linear regression modeling, the time constraint factor was the most powerful predictor of performance in adult learners, followed by the stimuli modality task factor, and when both task factors were considered together, they were highly predictive of performance with grammatical gender in Spanish as a nonnative language. The time constraint task factor was also predictive of performance in native speakers, although the nature of the time constraint effect was reversed for native speakers as they performed significantly better under a time constraint, contrary to their intermediate and beginner learner peers. Furthermore, task stimuli modality was *not* independently predictive of performance with grammatical gender in Spanish as a native language, contrary to the trend observed in adult learners.

In the following chapter, both the descriptive and inferential results will be presented for the effect of the individual factors on performance as well as the multilingual effect.

Chapter 5: Results: Individual factors and the multilingual effect

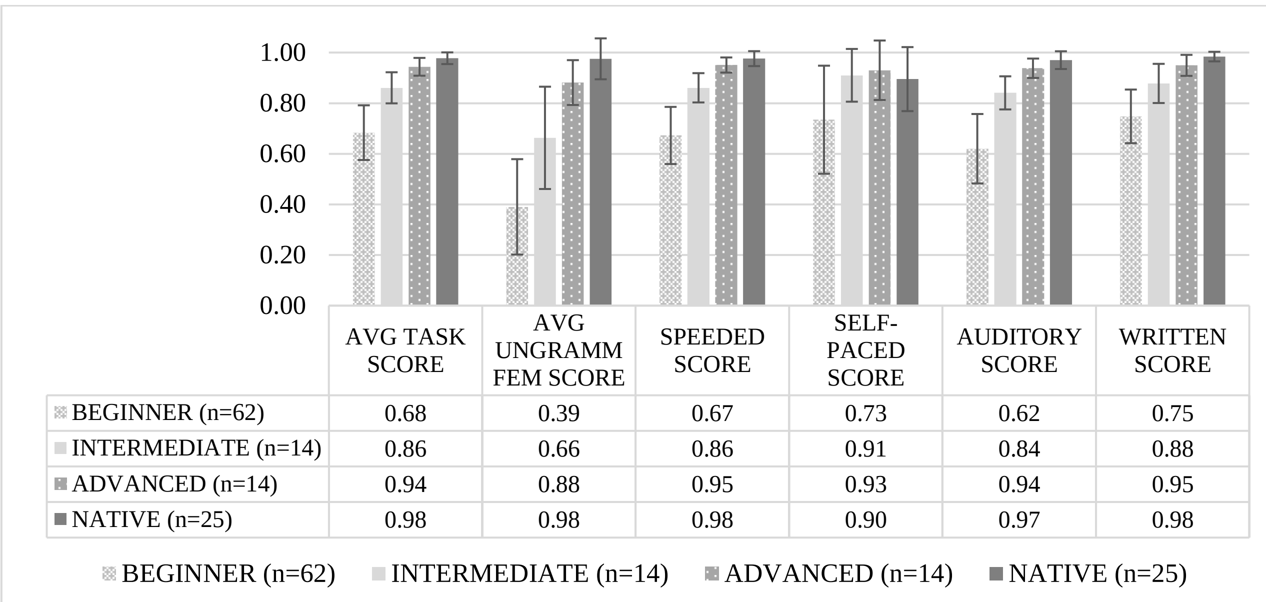
Recall that the data collected from the Language Learner Profile Questionnaire and the Spanish proficiency test (Task 1) are provided in the Methodology section. The results from each of the four experimental tasks (Task 2-5) and Task 6, the metalinguistic awareness exit survey, are presented below, according to the second two research questions posed concerning the relative effect of each individual factor on performance with grammatical gender in Spanish as well as the effect of being multilingual on learning an additional novel language.

5.1 Individual factors

The third research question asked how individual differences can predict or account for accuracy and variation in performance in adult learners of Spanish and further examined how these same individual differences may interact with task effects. In order to examine the effect of individual differences on accuracy and variation in performance, the individual factors were first examined categorically and adult learners were divided into groups according to the following independent variables (individual factors): Spanish proficiency level (beginner, intermediate, advanced), metacognitive awareness level (less vs. more aware), motivational orientation group (positive, negative, neutral), average weekly use of Spanish (low, moderate, high), metalinguistic awareness level (more vs. less aware), and Spanish attitude group (positive, negative, neutral). The corresponding accuracy scores and group variation (SD) of learner participants were then compared per individual factor group (as outlined above) across the different dependent response variables examined, including: average task score, average ungrammatical feminine score, average speeded score, average self-paced score, average auditory score, and average written score. It is important to note that by treating the independent/explanatory factors categorically, some unequal groupings emerged; nonetheless, these same factors were also treated as continuous variables for subsequent inferential analysis thereby mitigating this potential limitation of unequal categories. Finally, interactions between the individual factors are considered through inter-factor correlations in order to explore more closely the relationship and nature of the individual learner variables considered in this study. In what follows, the descriptive results of each individual factor treated as a categorical variable will be presented followed by the results of inferential analysis via correlations and linear regression modeling, treating each individual factor as a continuous variable, in order to identify any significant factors that may account for the variation observed in both native speakers and adult/late learners of Spanish.

5.1.1 Individual factors: Spanish proficiency

Regarding the effect of Spanish proficiency, self-reported and tested Spanish proficiency scores, including vocabulary, grammar, and total proficiency, are presented in the Methodology section along with a detailed description of how participants were ultimately categorized according to proficiency level. The descriptive results of the effect of tested Spanish proficiency (grammar and vocabulary scores combined) in both native speakers and adult learners are presented in Figure 17. Advanced learners (n = 14) clearly outperformed their other learner peers and their performance patterned closely with the native speaker group. The intermediate proficiency learners (n = 14) also clearly outperformed their beginner proficiency (n = 62) peers on all response variables examined. Clear distinctions between proficiency groups are most evident in average task scores, average scores on ungrammatical feminine noun targets, average speeded task scores, and on auditory stimuli scores; however, this clear distinction is less evident on average self-paced task scores (mean accuracy difference=2%) and on written scores (mean accuracy difference=7%), in which advanced and intermediate proficiency groups more closely resemble one another. Surprisingly, on average self-paced task scores, the advanced learner group slightly outperformed their native speaker peers by an average of 3%. The response variable that produced the greatest distinction between learner proficiency groups was their average scores on ungrammatical feminine noun targets (i.e., “AVG UNGRAMM FEM SCORE”), likely the most difficult target type, in which advanced learners clearly outperformed (M = .88, SD = .114) the intermediate learners (M = .66, SD = .200), who also clearly outperformed their beginner peers (M = .39, SD = .190). Spanish proficiency level also demonstrated a clear effect on group variation (SD) across tasks in which the advanced learner proficiency group also demonstrated the least variation among the adult learners that was also quite comparable to their native speaker peers, followed by the intermediate proficiency group whose variation was more than double that of their advanced proficiency peers on most tasks (with the exception of average self-paced task score), and the beginner proficiency group varied the most in their performance producing a standard deviation of nearly double that of their intermediate proficiency peers, as can be observed in the standard deviation bars shown in Figure 17. In summary, tested Spanish proficiency level appears to have an effect on both accuracy and variation across learner groups in which high proficiency learners pattern very closely to native Spanish speakers. Furthermore, higher proficiency level is associated with higher accuracy, particularly with ungrammatical feminine noun tokens and on auditory stimuli tasks, and variation in performance also appears to decrease with increasing proficiency level.



	AVG TASK SCORE	AVG UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
BEGINNER (n=62)	0.68	0.39	0.67	0.73	0.62	0.75
INTERMEDIATE (n=14)	0.86	0.66	0.86	0.91	0.84	0.88
ADVANCED (n=14)	0.94	0.88	0.95	0.93	0.94	0.95
NATIVE (n=25)	0.98	0.98	0.98	0.90	0.97	0.98

Figure 17. Learner proficiency groups and native speakers compared across average accuracy scores per task type. Standard deviation (SD) bars are provided.

The output of a Pearson's *r* correlation matrix between the different components of Spanish proficiency and accuracy scores on different task types is presented in Table 18 for adult learners and in Table 19 for native Spanish speakers. In the learner group, all four measures of Spanish proficiency, including self-report, vocabulary scores, grammar scores, and total combined tested proficiency, were significantly correlated with performance in terms of accuracy and intra-learner/speaker variance across the different task types. The strongest positive correlation was observed between tested vocabulary scores and speeded task scores ($r(89) = .803$, $p < .001$), followed by auditory task scores ($r(89) = .784$, $p < .001$), and ungrammatical feminine noun accuracy scores ($r(89) = .740$, $p < .001$). Overall, a strong, significant, and positive correlation was detected between total tested Spanish proficiency and average task scores ($r(89) = .765$, $p < .001$). However, it is interesting to note that between the subcomponents of vocabulary and grammar, stronger correlations were observed with the tested vocabulary score ($r(89) = .783$, $p < .001$) than with the tested grammar score ($r(89) = .632$, $p < .001$). Average intra-learner variation (expressed as variance) in learner accuracy scores demonstrated the strongest negative correlation with all components of Spanish proficiency, including total tested proficiency score ($r(89) = -.823$, $p < .001$), tested vocabulary score ($r(89) = -.827$, $p < .001$), tested grammar score ($r(89) = -.699$, $p < .001$), and self-reported Spanish proficiency ($r(89) = -.683$, $p < .001$). In summary, Spanish proficiency is significantly and positively correlated with performance in adult learners and this association is most pronounced on speeded and auditory tasks and seems to be more associated with tested vocabulary scores than tested grammar scores. Furthermore, intra-learner variance is highly and negatively correlated with all measures of Spanish proficiency, but again shows the strongest negative correlation with tested vocabulary scores in late learners.

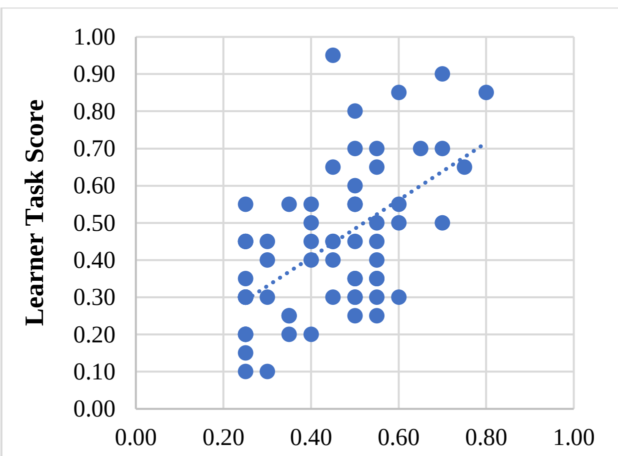
Table 18. Output of a correlation matrix (Pearson's *r*) examining the relationship between different measures of Spanish proficiency and accuracy and intra-learner variance on different task types in **learners**. Significant (95% CI) correlations are highlighted in grey.

	SPAN PROF: TOTAL SCORE		**SPAN PROF: SELF-REPORT**		**SPAN PROF: VOCAB SCORE**		**SPAN PROF: GRAMMAR SCORE**	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
AVG TASK SCORE	0.765	< .001	0.710	< .001	0.783	< .001	0.632	< .001
AVG VAR	-0.823	< .001	-0.683	< .001	-0.827	< .001	-0.699	< .001
AVG UNGRAM FEM SCORE	0.743	< .001	0.593	< .001	0.740	< .001	0.642	< .001
SPEEDED SCORE	0.782	< .001	0.748	< .001	0.803	< .001	0.644	< .001
SELF-PACED SCORE	0.415	< .001	0.333	0.001	0.408	< .001	0.365	< .001
AUDITORY SCORE	0.763	< .001	0.735	< .001	0.784	< .001	0.627	< .001
WRITTEN SCORE	0.653	< .001	0.570	< .001	0.662	< .001	0.545	< .001

In the native Spanish speaker group, significant correlations with Spanish proficiency were detected, although to a lesser extent than in the learner group. Table 19 presents the output of a Pearson's *r* correlation matrix between the different components of Spanish proficiency and the different measures of performance with grammatical gender, including average scores, intra-speaker variance, and accuracy scores on each of the task types in native Spanish speakers. A moderate and significant positive correlation was detected between tested vocabulary scores and average task scores ($r(24) = .441$, $p = .027$) in native Spanish speakers. However, contrary to the learners, native speakers demonstrated the strongest positive correlations with self-reported Spanish proficiency, as opposed to tested proficiency, and this effect was most pronounced for average scores on ungrammatical feminine noun tokens ($r(24) = .767$, $p < .001$), followed by speeded scores ($r(24) = .663$, $p < .001$), and finally auditory scores ($r(24) = .590$, $p = .002$). In terms of correlations with total tested Spanish proficiency, the only significant correlation detected was between average ungrammatical feminine noun scores and total Spanish proficiency ($r(24) = .396$, $p = .050$), however, no other significant correlations were detected with total tested Spanish proficiency. Most notably, the native speaker group demonstrated no significant correlations between their tested grammar scores and any of their task scores, in stark contrast to the learner group who demonstrated consistently strong correlations between tested grammar scores and their average scores on all task types (see Table 18). In summary, native Spanish speakers demonstrated some moderate to strong correlations between their Spanish proficiency scores and scores on the different task types. However, correlations were larger and more significant with self-reported Spanish proficiency than with tested Spanish proficiency and notably no significant correlations were detected between tested grammar scores and task scores in native speakers. Figure 18 demonstrates the more consistent and stronger correlation observed between Spanish proficiency and task scores in adult learners as compared to their native speaker counterparts.

Table 19. Output of a correlation matrix (Pearson's *r*) examining the relationship between different measures of Spanish proficiency and accuracy and intra-speaker variance on different task types in **native Spanish speakers**. Significant (95% CI) correlations are highlighted in grey.

	SPAN PROF: TOTAL SCORE		SPAN PROF: SELF-REPORT		SPAN PROF: VOCAB SCORE		SPAN PROF: GRAMMAR SCORE	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*
AVG TASK SCORE	0.308	0.134	0.550	0.004	0.441	0.027	0.136	0.518
AVG VAR	-0.290	0.160	-0.539	0.005	-0.417	0.038	-0.124	0.555
AVG UNGRAM FEM SCORE	0.396	0.050	0.767	< .001	0.559	0.004	0.189	0.364
SPEEDED SCORE	0.314	0.126	0.663	< .001	0.489	0.013	0.107	0.611
SELF-PACED SCORE	0.364	0.074	0.420	0.037	0.309	0.132	0.369	0.070
AUDITORY SCORE	0.242	0.243	0.590	0.002	0.404	0.045	0.058	0.781
WRITTEN SCORE	0.270	0.191	0.321	0.118	0.348	0.088	0.149	0.477



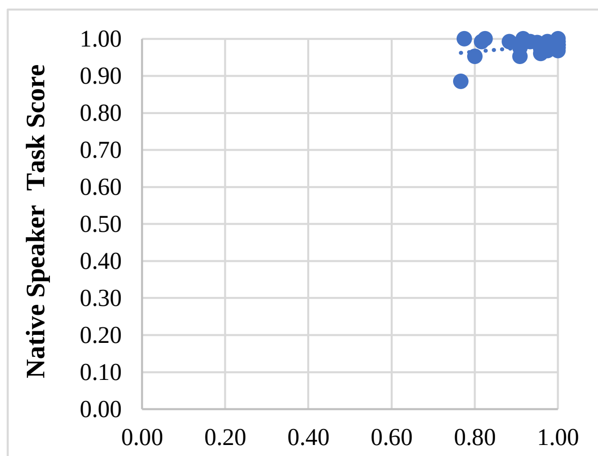
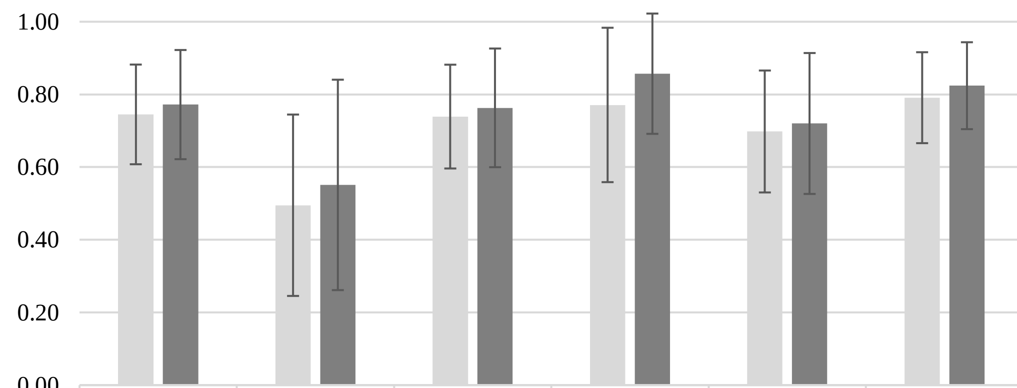


Figure 18. Correlations compared between total Spanish proficiency and average task scores in **learners (left)** and **native Spanish speakers (right)**.

5.1.2 Individual factors: metacognitive awareness

In order to examine the effect of metacognitive awareness on learner accuracy scores and group variation (SD), adult learner participants were categorized into two main groups based on their quantified responses to the metacognitive awareness section of the Language Learner Profile Questionnaire: "less aware" (score < 3) or "more aware" (score > 3). The descriptive results of this analysis are presented in Figure 19. More metacognitively aware learners outperformed their less metacognitively aware peers on every measure, although the degree of difference varied according to response variable; in terms of average task score, *more aware* learners (n = 22) demonstrated an advantage of only 2% over their *less aware* learner peers (n = 68). However, the largest advantage for more metacognitively aware learners was observed on their average self-paced task score (9% higher), followed by their average scores on ungrammatical feminine nouns (5% higher), and on their average written task scores (3% higher). In terms of group variation (SD) in learner performance, metacognitively *more aware* and *less aware* learners were quite comparable, as demonstrated in the very similar standard deviation bars shown in Figure 19. In summary, more metacognitively aware learners appear to demonstrate higher accuracy scores, and this advantage is particularly evident during self-paced tasks and on the more difficult ungrammatical feminine noun token type. However, being more metacognitively aware does not appear to be associated with lower variation in scores.



	AVG TASK SCORE	AVG UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
LESS AWARE (n=68)	0.75	0.50	0.74	0.77	0.70	0.79
MORE AWARE (n=22)	0.77	0.55	0.76	0.86	0.72	0.82

LESS AWARE (n=68) MORE AWARE (n=22)

Figure 19. Metacognitive awareness groups compared across accuracy scores in **adult learners**. Standard deviation (SD) bars are provided.

Metacognitive awareness level and type were also examined as continuous variables in adult Spanish learners. The output of a Pearson's *r* correlation matrix between the different components of metacognitive awareness and average scores on the different task types is presented in Table 20. Learners demonstrated moderate significant correlations between the *metacognitive knowledge* component and average auditory scores (*r*(89) = .232, *p* = .028) as well as average speeded scores (*r*(89) = .220, *p* = .038). In addition, a significant correlation was detected between learners' average auditory task scores (*r*(89) = .208, *p* = .049) and their total metacognitive awareness scores. However, no significant correlations were detected with the *regulation* component of metacognitive awareness nor with average task scores. In summary, metacognitive awareness appears to have a moderate effect on learners' accuracy, an effect particularly associated with speeded and auditory tasks. Nonetheless, the effect of metacognitive awareness appears to be limited to *knowledge about* cognition and not the *regulation of* cognition and is only detectable with certain task types (speeded and auditory) but is not significantly correlated with task scores overall. Figure 20 demonstrates the moderate, yet non-significant correlation observed between total metacognitive awareness and average task scores in adult learners.

Table 20. Output of a correlation matrix (Pearson's *r*) examining the relationship between the different dimensions of metacognitive awareness (knowledge about and regulation of cognition) and accuracy and intra-learner variance on different task types in **learners** of Spanish. Significant (95% CI) correlations are highlighted in grey.

	METACOG: KNOWLEDGE		METACOG: REGULATION		METACOG SCORE	
	r	*p*	*r*	*p*	*r*	*p*
AVG TASK SCORE	0.190	0.072	0.102	0.341	0.174	0.100
AVG VAR AVG UNGRAM	-0.151	0.157	-0.089	0.402	-0.143	0.179
FEM SCORE	0.155	0.144	0.068	0.526	0.135	0.206
SPEEDED SCORE	0.220	0.038	0.112	0.295	0.199	0.060
SELF-PACED SCORE	0.130	0.221	0.119	0.262	0.144	0.174
AUDITORY SCORE	0.232	0.028	0.114	0.285	0.208	0.049
WRITTEN SCORE	0.105	0.324	0.067	0.528	0.101	0.343

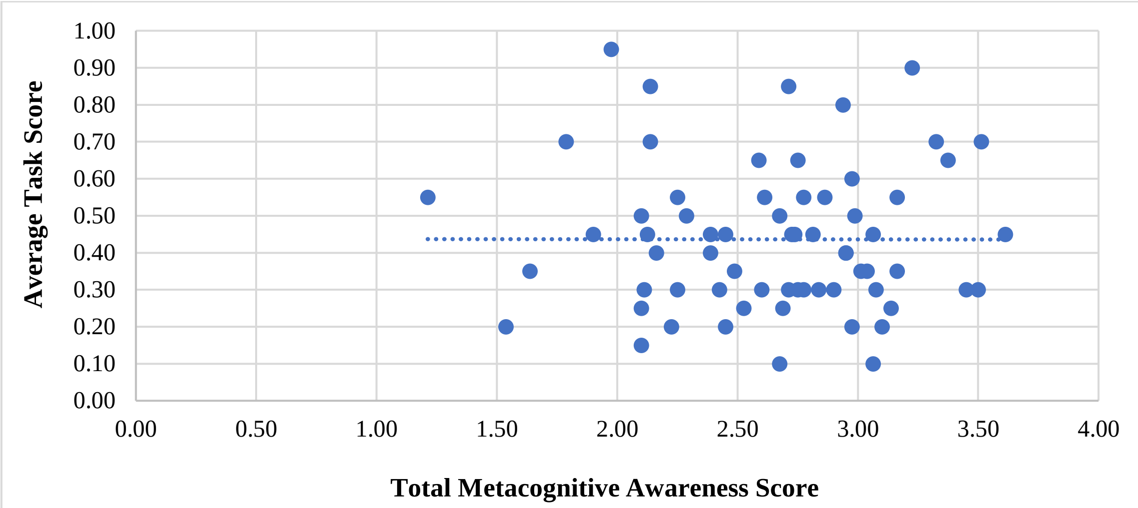
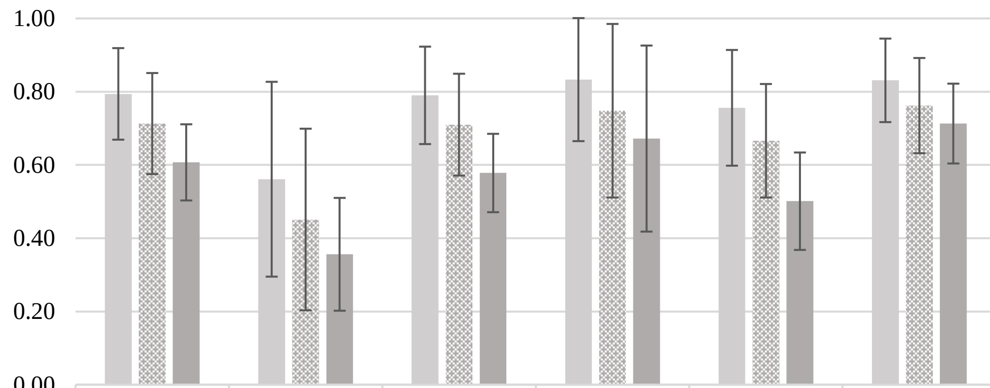


Figure 20. Correlation between average task score and total metacognitive awareness score in **learners** only.

5.1.3 Individual factors: motivational orientation

The effect of motivation level on accuracy scores and intra-learner variance and group variation (SD) were investigated by dividing learners into three groups based on their responses to the motivational orientation section of the Language Learner Profile Questionnaire on a scale of *totally disagree* (-2) to *totally agree* (+2): "positive" (scores > 1), "negative" (scores < -1), and "neutral" (-1 < scores > 1). The descriptive results of this analysis are provided in Figure 21. Clear distinctions between motivational orientation groups can be observed in which learners with a *positive* motivational orientation clearly outperformed their peers on every measure. Although the *positive* (n = 56) and *neutral* (n = 24) motivational orientation groups tended to resemble each other across all response variables, clearer distinctions emerged between the *positive* motivational orientation group and the *negative* motivational orientation group (n = 10). Most notably, on average auditory task scores, the *positive* motivational orientation group vastly outperformed (M = .76, SD = .16) their *negative* motivational orientation peers (M = .50, SD = .12), resulting in an average auditory task score difference of 26%. This same trend was also observed on average speeded task scores, resulting in an 21% advantage for the *positive* motivational orientation group, and on average scores on ungrammatical feminine nouns, resulting in a 20% advantage. With regards to variation as a group (SD) in learner performance, the three motivational groups appear to be quite similar. In fact, surprisingly, the *negative* motivational orientation group demonstrated less variation in their scores than their peers on all measures except average self-paced task scores. In summary, adult learners with a more positive motivational orientation consistently demonstrate higher accuracy than their peers with a more negative motivational orientation, and this advantage is particularly evident on speeded and auditory tasks. Nonetheless, this distinction appears to be more binary in nature in which the clearest distinctions are observable between positive and negative motivational orientations and the effect of motivational orientation on variation in learner performance is less evident.



	AVG TASK SCORE	AVG UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
POSITIVE (n=56)	0.79	0.56	0.79	0.83	0.76	0.83
NEUTRAL (n=24)	0.71	0.45	0.71	0.75	0.67	0.76
NEGATIVE (n=10)	0.61	0.36	0.58	0.67	0.50	0.71

Figure 21. Motivational orientation groups compared across accuracy scores in **learners**. Standard deviation (SD) bars are provided.

Motivation type and level were also examined as continuous variables in adult Spanish learners. The output of a Pearson's *r* correlation matrix between the different components of motivation and average scores on the different task types is presented in Table 21. Significant moderate to strong correlations between average task scores and motivation were detected with the *ideal L2 self* component of motivation ($r(89) = .520$, $p < .001$), total motivation scores ($r(89) = .498$, $p < .001$), the *experience* component of motivation ($r(89) = .341$, $p < .001$), and finally with the *ought-to L2 self* component of motivation ($r(89) = .326$, $p = .002$). However, no significant correlations were detected with the behavior component of motivation on average task scores nor average scores on any task type. Furthermore, the strongest significant correlations were detected between the *ideal L2 self* component and average speeded task scores ($r(89) = .562$, $p < .001$), average auditory task scores ($r(89) = .562$, $p < .001$), as well as between *total motivational score* and average speeded task scores ($r(89) = .516$, $p < .001$) and average auditory task scores ($r(89) = .509$, $p < .001$). With regards to variation, significant negative correlations were detected between average variation and the *ideal L2 self* component ($r(89) = -.448$, $p < .001$), *total motivational score* ($r(89) = -.434$, $p < .001$), the *ought-to L2 self* component ($r(89) = -.333$, $p = .001$), and the *experience* component of motivation ($r(89) = -.291$, $p = .005$). In summary, motivational orientation had a moderate to strong positive correlation with accuracy scores across tasks and was also negatively correlated with intra-learner variance; however, this association with motivation was most pronounced with speeded and auditory tasks. Furthermore, the *ideal L2 self*, the *ought-to L2 self*, the *experiential* component of motivation, as well as *overall motivation* scores were all significantly and positively correlated with accuracy scores, while the *behavioral* component of motivation was not correlated with performance on any task in adult learners of Spanish. Figure 22 demonstrates the positive correlation observed between learners' *total motivational scores* and average task scores while Figure 23 demonstrates the correlation between the *ideal L2 self* and average task scores.

Table 21. Output of a correlation matrix (Pearson's *r*) examining the relationship between motivational orientation (motivated learning behavior, the ideal L2 self, the ought-to L2 self, motivated learning experience, and total score) and accuracy and intra-learner variance on different task types in **learners**. Significant (95% CI) correlations are highlighted in grey.

	MOTIV: BEHAVIOR		MOTIV: IDEAL L2 SELF		MOTIV: OUGHT-TO L2 SELF		MOTIV: EXPERIENCE		MOTIV: TOTAL SCORE	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
AVG TASK SCORE	0.133	0.211	0.520	< .001	0.326	0.002	0.341	0.001	0.498	< .001
AVG VAR	-0.055	0.605	-0.448	< .001	-0.333	0.001	-0.291	0.005	-0.434	< .001
AVG UNGRAM FEM SCORE	-0.002	0.985	0.368	< .001	0.287	0.006	0.213	0.044	0.339	0.001
SPEEDED SCORE	0.121	0.254	0.562	< .001	0.342	< .001	0.339	0.001	0.516	< .001
SELF-PACED SCORE	0.194	0.066	0.204	0.054	0.101	0.342	0.305	0.003	0.282	0.007
AUDITORY SCORE	0.117	0.271	0.562	< .001	0.332	0.001	0.333	0.001	0.509	< .001
WRITTEN SCORE	0.136	0.202	0.386	< .001	0.258	0.014	0.303	0.004	0.404	< .001

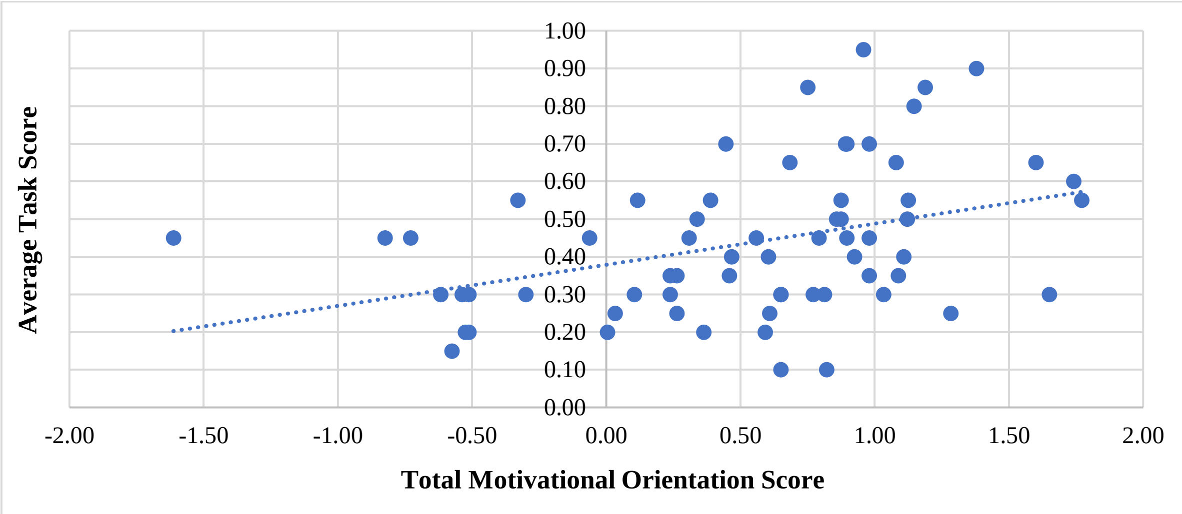


Figure 22. Correlation between average task score and total motivational score in **learners** only.

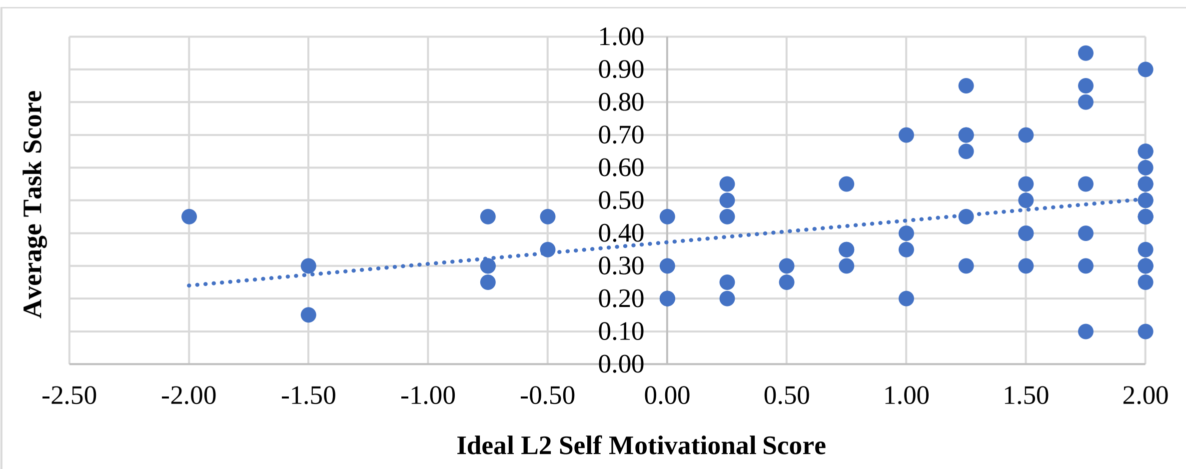
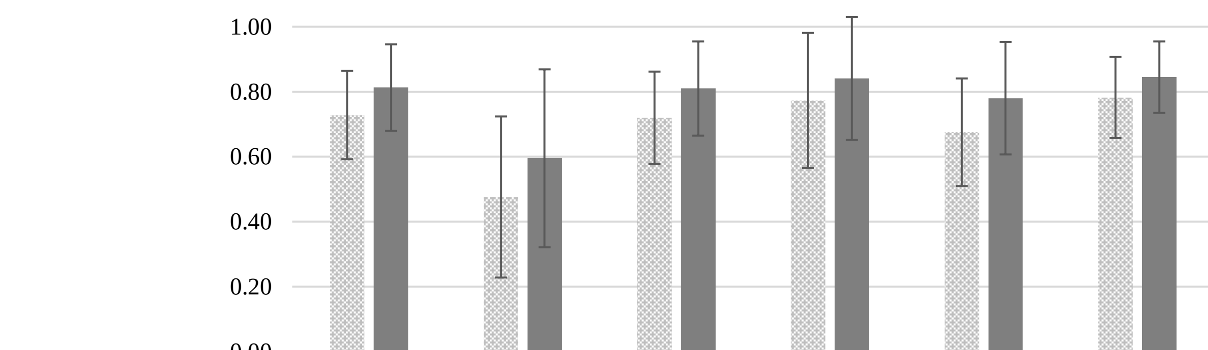


Figure 23. Correlation between average task score and ideal L2 self-motivational score in **learners** only.

5.1.4 Individual factors: Spanish use

The effect of Spanish language use was also examined in both native Spanish speakers and adult learner groups. After being divided according to *speaker status*, that is *native speaker* versus *adult learner*, both participant groups were further categorized into subgroups based on their reported use of Spanish averaged across five different contexts (with friends, with family, at school/work, when engaging in self-talk, and when counting) during a typical week. The majority of adult learners were categorized in the "low use" group (n=65), reporting an average weekly use of Spanish between 0% and 10%. The rest of the learners were categorized into the “moderate use” group (n=25), reporting an average weekly use of Spanish between 11% and 50%. The descriptive results of the effect of average Spanish use in learners are presented in Figure 24. The *moderate use* learner group consistently outperformed their *low use* peers on all measures analyzed, demonstrating the greatest advantage on accuracy with ungrammatical feminine nouns, resulting in a 12% difference in accuracy scores between *moderate use* learners (M = .60, SD = .27) and *low use* learners (M = .48, SD = .25). Similarly, *moderate use* learners (M = .78, SD = .17) outperformed their *low use* peers (M = .68, SD = .17) on auditory stimuli tasks by an average of 10%. Overall, *moderate use* learners demonstrated a clear advantage on average task scores (M = .81, SD = .13) compared to their *low use* peers (M = .73, SD = .14). With regards to group variation, both learner groups were remarkably similar, as demonstrated in the error (SD) bars in Figure 24. In summary, adult learners who reported moderate use of Spanish in a typical week clearly outperformed their peers with low reported use of Spanish, particularly with the more difficult token type of ungrammatical feminine nouns and on auditory tasks.



	AVG TASK SCORE	UNGRAM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
LOW USE (n=65)	0.73	0.48	0.72	0.77	0.68	0.78
MODERATE USE (n=25)	0.81	0.60	0.81	0.84	0.78	0.85

Figure 24. Spanish language use groups compared across accuracy scores in **learners**. Standard deviation (SD) bars are provided.

The majority of native Spanish speakers were categorized in the “high use” group (n=15), reporting average use of Spanish in a typical week of over 50% and the rest of the native Spanish speakers fell within the “moderate use” group (n=10), reporting average use of Spanish in a typical week of between 11% and 50%. The descriptive results of the effect of average Spanish use in native Spanish speakers are presented in Figure 25. For the most part, differences in average Spanish language use in a typical week did not lead to differences in performance with grammatical gender in native Spanish speakers, with two notable exceptions: *high use* native speakers seemed to have outperformed their *moderate use* peers on ungrammatical feminine noun tokens (*high use*: M = .99, SD = .02; *low use*: M = .95, SD = .12) and on average accuracy scores on the self-paced tasks (*high use*: M = 91, SD = .13; *low use*: M = 87, SD = .13), producing a 4% advantage for *high use* native speakers on both measures. With regards to variation, reported Spanish language use also appears to have an effect such that higher average weekly use of Spanish is associated with lower group variation (SD) in performance, with the exception of the self-paced tasks in which both native speaker groups demonstrate remarkably similar variation, as shown in the error (SD) bars in Figure 25. In summary, native Spanish speaker performance seems to be largely unaffected by average use of Spanish; however, some advantage is observed in performance with ungrammatical feminine noun tokens and on self-paced tasks for those native speakers who report high use of Spanish on a weekly basis.

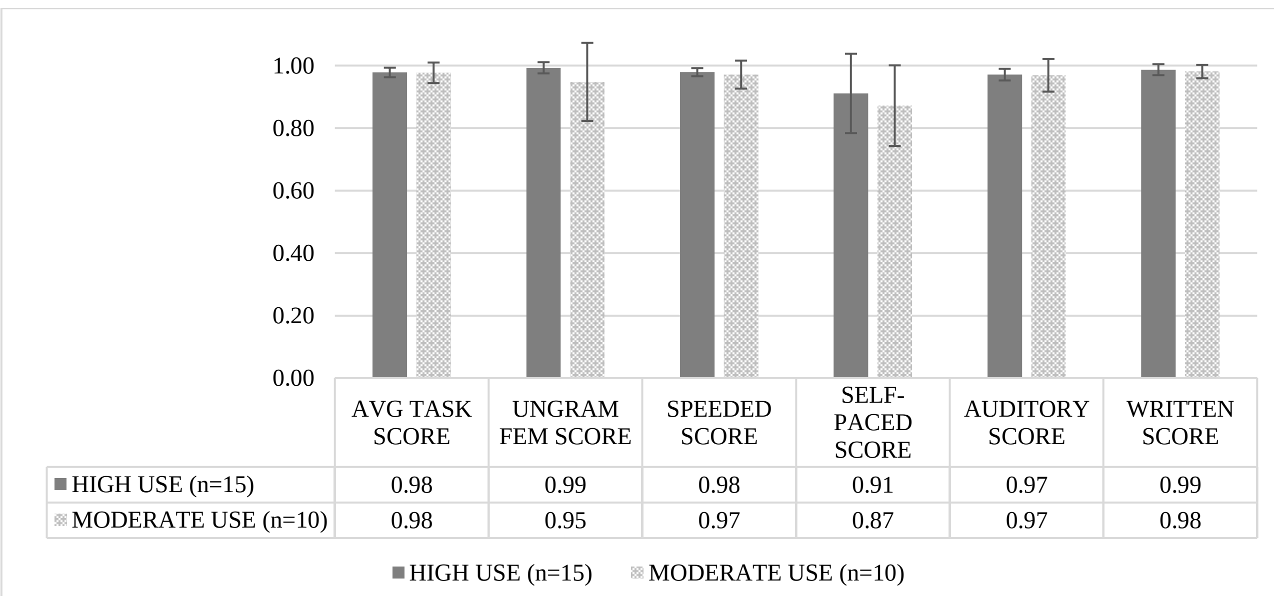


Figure 25. Spanish language use groups compared across accuracy scores in **native Spanish speakers**. Standard deviation (SD) bars are provided.

Spanish language use was also examined as a continuous variable in both native Spanish speakers and adult Spanish learners. The output of a Pearson's r correlation matrix between self-reported weekly Spanish use and average scores on the different task types is presented in Table 22 for adult learners and in Table 23 for native Spanish speakers. In the learner group, significant positive and moderate correlations were detected between average weekly reported use of Spanish and average task scores (*r*(89) = .288, *p* = .006), speeded scores (*r*(89) = .290, *p* = .006), auditory scores (*r*(89) = .278, *p* = .008), and written scores (*r*(89) = .259, *p* = .014). A significant negative correlation was detected between learners' average weekly reported use of Spanish and average intra-learner variance (*r*(89) = -.295, *p* = .005). Interestingly, learners' average reported weekly use of Spanish was not found to be significantly correlated with performance on self-paced tasks nor with the more difficult token type of ungrammatical feminine nouns. In summary, learners who use Spanish more often in their daily lives across different contexts tend to have moderately higher accuracy in their performance with grammatical gender on auditory, written, and speeded tasks, and tend to demonstrate less variation (expressed as intra-learner variance) in their performance, yet this benefit of language use does not seem to be associated with tasks in which learners can take their time (i.e., self-paced tasks) nor does it seemed to be linked to their performance with more difficult token types (i.e., ungrammatical feminine nouns).

Table 22. Output of a correlation matrix (Pearson's *r*) examining the relationship between self-reported weekly use of Spanish (averaged across five different daily life contexts) and accuracy and intra-learner variance on different task types in **learners**. Significant (95% CI) correlations are highlighted in grey.

	SPAN USE SCORE	
	r	*p*
AVG TASK SCORE	0.288	0.006
AVG VAR	-0.295	0.005
AVG UNGRAM FEM SCORE	0.182	0.086
SPEEDED SCORE	0.290	0.006
SELF-PACED SCORE	0.159	0.135
AUDITORY SCORE	0.278	0.008
WRITTEN SCORE	0.259	0.014

In native speakers of Spanish, a significant, positive and moderate correlation was detected between average reported weekly use of Spanish and average accuracy scores on ungrammatical feminine nouns only (*r*(24) = .411, *p* = .041), as shown in Table 23. No other significant correlations were detected with average Spanish use in native speakers.

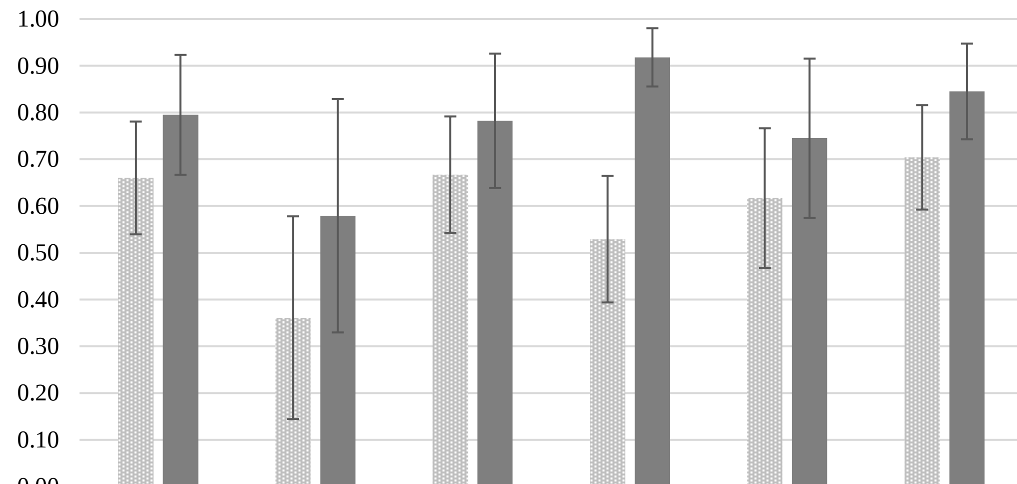
Table 23. Output of a correlation matrix (Pearson's *r*) examining the relationship between self-reported weekly use of Spanish (averaged across five different daily life contexts) and accuracy and intra-speaker variance on different task types in **native Spanish speakers**. Significant (95% CI) correlations are highlighted in grey.

	SPAN USE SCORE	
	r	*p*
AVG TASK SCORE	0.103	0.623
AVG VAR	-0.107	0.610
AVG UNGRAM FEM SCORE	0.411	0.041
SPEEDED SCORE	0.264	0.202
SELF-PACED SCORE	0.103	0.623
AUDITORY SCORE	0.139	0.509
WRITTEN SCORE	0.169	0.418

5.1.5 Individual factors: metalinguistic awareness

The effect of metalinguistic awareness was also examined in both native Spanish speakers and adult learner groups. After being divided according to *speaker status*, that is, *native speaker* versus *adult learner*, both participant groups were further categorized into two groups based on how they responded to the metalinguistic awareness exit survey (Task 6), producing an average score of their overt and conditioned metalinguistic awareness, both treated as a binary (1-0) in which each participant was either overtly aware (1) or not (0) of the grammatical gender structure being tested, and if they were not overtly aware, if that awareness could be conditioned by the researcher during the exit survey via examples of common grammatical gender errors and probing follow-up questions (see: *Methodology* for further explanation). This dichotomy produced two categories: "more aware" (mean metalinguistic awareness score = 1) and "less aware" (mean metalinguistic awareness score < 1). The descriptive results of this analysis are presented in Figure 27 for adult learners and in Figure 28 for native Spanish speakers.

Among the learner participants, *more aware* individuals (n = 61) outperformed their *less aware* peers (n = 29) on every measure. As can be observed in Figure 26, this distinction between metalinguistic awareness learner groups was most pronounced on self-paced task scores, producing a 39% difference in accuracy scores, as well as on average scores on ungrammatical feminine nouns, producing a 22% difference, and finally on average written task scores, producing a difference of 15%. Furthermore, *more aware* learners also demonstrated less group variation (SD) in their accuracy scores than their *less aware* peers, particularly on their average self-paced task scores. In other words, when given sufficient time to self-pace their performance, more metacognitively aware learners vary much less in their accuracy than their less metacognitively aware learner peers. Therefore, it is clear that being more explicitly aware of the grammatical structure being tested (i.e., metalinguistic awareness) affords certain advantage in adult learners, particularly on self-paced and written tasks and on the more difficult token type of ungrammatical feminine target nouns, and a higher level of metalinguistic awareness also appears to be associated with lower variation in learner performance.

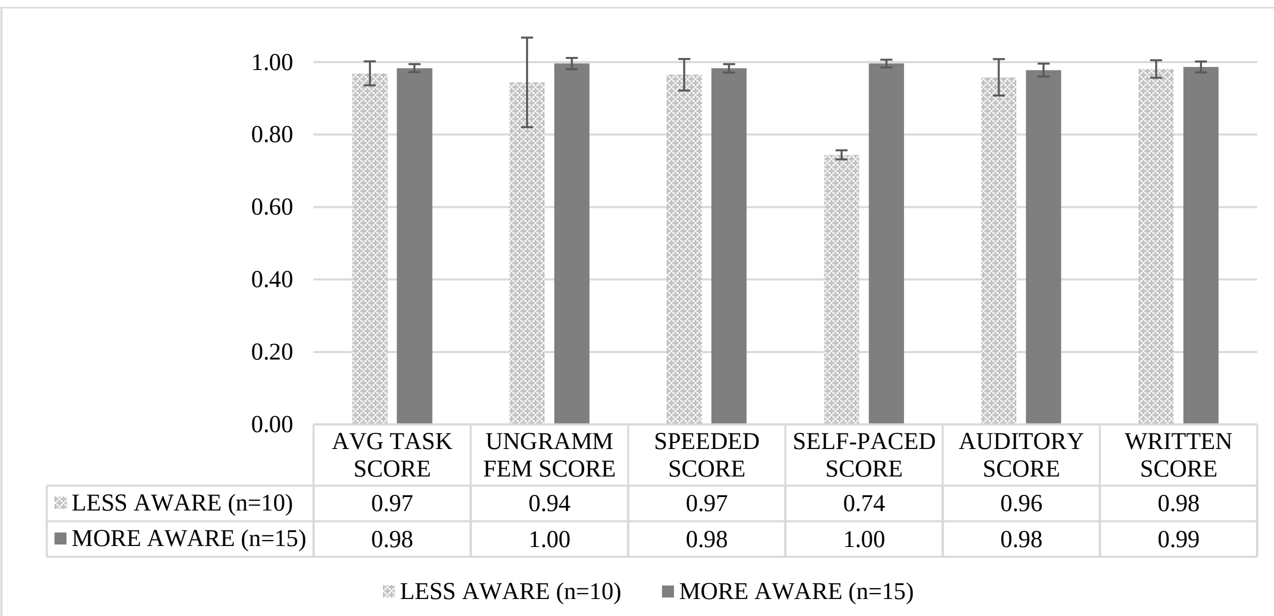


	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
LESS AWARE (n=29)	0.66	0.36	0.67	0.53	0.62	0.70
MORE AWARE (n=61)	0.80	0.58	0.78	0.92	0.75	0.85

LESS AWARE (n=29) MORE AWARE (n=61)

Figure 26. Metalinguistic awareness groups compared across accuracy scores in **learners** only. Standard deviation (SD) bars are provided.

Surprisingly, native Spanish speakers also appear to exhibit an effect of their metalinguistic awareness level. As demonstrated in Figure 27, *more aware* (n = 15) native speakers outperformed their *less aware* peers (n = 10) on every measure, although the difference in accuracy scores between the two groups was relatively small compared to the learner group. The effect of metalinguistic awareness on native speaker performance was most evident on average self-paced task scores, resulting in a 26% difference between *more aware* (*M* = .996, *SD* = .011) and *less aware* (*M* = .744, *SD* = .013) native Spanish speaker participants. In addition, *more aware* native speakers performed better (*M* = .996, *SD* = .016) with the more difficult token type of ungrammatical feminine nouns, producing an accuracy advantage of 6% compared to their *less aware* peers (*M* = .944, *SD* = .124). Regarding group variation (SD) in average scores, *more aware* native speakers on average demonstrated less variation in their performance than those native speakers who were *less* metalinguistically *aware* of the grammatical gender structure being tested, as can be observed in the respective error bars per group represented in Figure 27. Nonetheless, the average task score difference between more and less aware native speakers was rather minimal compared to that of their adult learner counterparts, resulting in an average task score advantage for *more aware* native speakers of only 1%. In summary, native speakers' accuracy on tasks was also affected by their metalinguistic awareness of grammatical gender, yet to a lesser extent than their learner counterparts, and this effect in native speakers was most pronounced on self-paced tasks and on the more difficult token type of ungrammatical feminine target nouns and further appears to be associated with relatively less variation in performance.



	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
LESS AWARE (n=10)	0.97	0.94	0.97	0.74	0.96	0.98
MORE AWARE (n=15)	0.98	1.00	0.98	1.00	0.98	0.99

Figure 27. Metalinguistic awareness groups compared across accuracy scores in **native Spanish speakers** only. Standard deviation (SD) bars are provided.

The output of a Pearson's *r* correlation matrix between the different forms of metalinguistic awareness and average accuracy scores on each task type is presented in Table 24 for adult Spanish learners and in Table 25 for native Spanish speakers. The learners demonstrated by far the strongest significant positive correlation between *total metalinguistic awareness* and average accuracy on self-paced tasks (*r*(89) = .948, *p* < .001), followed by a significant positive strong correlation between *overt metalinguistic awareness* and average accuracy on self-paced tasks (*r*(89) = .894, *p* < .001), and *conditioned metalinguistic awareness* (*r*(89) = .611, *p* < .001). In addition, average accuracy scores on written tasks were also significantly and positively correlated with both *total metacognitive awareness* (*r*(89) = .561, *p* < .001) and *overt metalinguistic awareness* (*r*(89) = .533, *p* < .001), as well as to a lesser extent with *conditioned metalinguistic awareness* (*r*(89) = .353, *p* < .001). With regards to overt compared to conditioned metalinguistic awareness, as can be observed in Table 24, *conditioned* awareness produced weaker and, on average, less significant correlations with accuracy than did *overt* and *total* metalinguistic awareness. Furthermore, with regards to intra-learner variance, significant negative correlations were detected with *total* metalinguistic awareness (*r*(89) = -.450, *p* < .001), *overt* metalinguistic awareness (*r*(89) = -.427, *p* < .001), and, to a lesser extent, with *conditioned* metalinguistic awareness (*r*(89) = -.284, *p* = .007). Overall, learners demonstrated moderate to strong significant positive correlations between their accuracy scores on all tasks and their metalinguistic awareness scores, and this correlation was particularly pronounced with self-paced and written tasks. Furthermore, learners' metalinguistic awareness level was moderately and negatively correlated with their average intra-learner variance.

Table 24. Output of a correlation matrix (Pearson's *r*) examining the relationship between metalinguistic awareness (overt, conditioned, total) and accuracy and intra-learner variance on different task types in **learners**. Significant (95% CI) correlations are highlighted in grey.

	METALING AWARE: OVERT		**METALING AWARE: CONDITIONED**		**METALING AWARE: TOTAL**	
	r	*p*	*r*	*p*	*r*	*p*
AVG TASK SCORE	0.451	< .001	0.333	0.001	0.488	< .001
AVG VAR	-0.427	< .001	-0.284	0.007	-0.450	< .001
AVG UNGRAM FEM SCORE	0.395	< .001	0.247	0.019	0.410	< .001
SPEEDED SCORE	0.366	< .001	0.306	0.003	0.409	< .001
SELF-PACED SCORE	0.894	< .001	0.611	< .001	0.948	< .001
AUDITORY SCORE	0.346	< .001	0.281	0.007	0.384	< .001
WRITTEN SCORE	0.533	< .001	0.353	< .001	0.561	< .001

Native Spanish speakers also showed some significant correlations between their accuracy scores and metalinguistic awareness, yet overall, to a somewhat lesser extent than the learner group, as shown in Table 25. Native speakers demonstrated an equally strong positive correlation between average accuracy scores on self-paced tasks and *overt metalinguistic awareness* ($r(24) = .996$, $p < .001$) as well as *total metalinguistic awareness* ($r(24) = .996$, $p < .001$). Furthermore, an equally strong positive correlation was detected between *conditioned metalinguistic awareness* and average scores on ungrammatical feminine nouns ($r(24) = .996$, $p < .001$). Finally, a strong correlation was detected between *conditioned metalinguistic awareness* and average scores on auditory tasks ($r(24) = .831$, $p < .001$) in the native Spanish speaker group. Contrary to their adult learner counterparts, native Spanish speakers demonstrated stronger correlations with *conditioned* metalinguistic awareness than with *overt* awareness and *total* metalinguistic awareness. Also contrary to the learner group, native speakers demonstrated no significant correlations between either type of metalinguistic awareness and average intra-speaker variance. In summary, native speaker accuracy scores were correlated with their level of metalinguistic awareness about grammatical gender as the target structure, yet to a lesser extent than their adult learner peers. Furthermore, this correlation was most pronounced with conditioned awareness rather than overt or total metalinguistic awareness, and average intra-speaker variance was *not* correlated with metalinguistic awareness in native Spanish speakers.

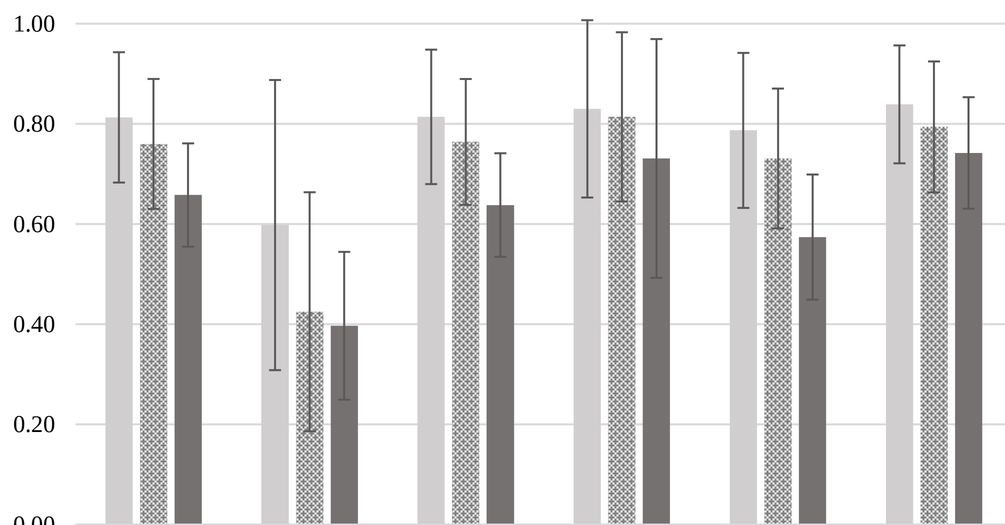
Table 25. Output of a correlation matrix (Pearson's *r*) examining the relationship between metalinguistic awareness (overt, conditioned, total) and accuracy and intra-speaker variance on different task types in **native Spanish speakers**. Significant (95% CI) correlations are highlighted in grey.

	METALING AWARE: OVERT		**METALING AWARE: CONDITIONED**		**METALING AWARE: TOTAL**	
	r	*p*	*r*	*p*	*r*	*p*
AVG TASK SCORE	0.310	0.132	0.789	0.133	0.310	0.132
AVG VAR	-0.292	0.157	-0.772	0.154	-0.292	0.157
AVG UNGRAM FEM SCORE	0.321	0.117	0.966	< .001	0.321	0.117
SPEEDED SCORE	0.300	0.145	0.888	< .001	0.300	0.145
SELF-PACED SCORE	0.996	< .001	0.239	0.250	0.996	< .001
AUDITORY SCORE	0.285	0.168	0.831	< .001	0.285	0.168
WRITTEN SCORE	0.150	0.476	0.488	0.013	0.150	0.476

5.1.6 Individual factors: attitudes

Although all native Spanish speakers responded positively to the Spanish language attitudes portion of the Language Learner Profile Questionnaire, varied attitudes were observed in the learner group. In order to investigate the possible effect of attitudes towards the Spanish language and target language community, adult learner participants were subdivided according to their responses on the attitudes section of the Language Learner Profile Questionnaire, along a scale of *strongly disagree* (-2) to *strongly agree* (+2), and were categorized into three groups: "positive" (scores > 1), "negative" (scores < -1), and "neutral" (-1 < scores > 1). The descriptive results of this analysis are presented in Figure 28. A positive and robust effect of attitude was observed among the learner participants in which those learners with *positive attitudes* towards the target language and community (n = 49) consistently outperformed their *negative attitude* (n = 33) and *neutral attitude* (n = 8) peers. On average, *positive attitude* learners scored 15% higher (M = .813, SD = .130) than their *negative attitude* peers (M = .658, SD = .103). The difference in average accuracy scores per attitude group was most pronounced on average auditory task scores, producing a 22% difference in mean accuracy, average speeded task scores, producing a 17% difference in mean accuracy, and on the ungrammatical feminine noun tokens, producing a 20%

difference in accuracy scores, as observed in Figure 28. With regards to group variation (SD) in their accuracy scores, the attitudinal learner groups were quite comparable and *positive attitude* learners actually demonstrated relatively more variation than their peers on the ungrammatical feminine noun tokens, as demonstrated in the error bars per attitude group in Figure 28. Furthermore, although a clear distinction in average accuracy scores is evident between *positive* and *negative* attitude learner groups, this distinction is far less pronounced between the *positive* and *neutral* attitude learner groups and the least difference between learner attitude groups is observed on the written and self-paced tasks. In summary, adult learners with more positive attitudes toward the target language and target language community demonstrated consistently higher accuracy rates than their negative attitude peers, and this effect was particularly pronounced during the speeded and auditory tasks as well as with the more difficult ungrammatical feminine noun tokens. Nonetheless, similarly to the effect of motivation, the effect of attitudes seems to be more binary in nature in which the clearest distinctions are observable between positive and negative attitude groups, yet the effect of a more neutral attitude is less clear.



	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
POSITIVE (n=49)	0.81	0.60	0.81	0.83	0.79	0.84
NEUTRAL (n=8)	0.76	0.43	0.76	0.81	0.73	0.79
NEGATIVE (n=33)	0.66	0.40	0.64	0.73	0.57	0.74

Figure 28. Spanish attitude groups compared across accuracy scores in **learners** only. Standard deviation (SD) bars are provided.

The output of a Pearson's *r* correlation matrix between average Spanish attitude scores and average accuracy scores on each task type is presented in Table 26. Spanish attitude scores were significantly correlated with all measures in the learner group with the exception of self-paced accuracy scores. Average task scores were moderately and positively correlated with Spanish attitudes scores ($r(89) = .484$, $p < .001$). However, the correlation with Spanish attitudes was most pronounced on auditory task scores ($r(89) = .552$, $p < .001$) and on speeded task scores ($r(89) = .538$, $p < .001$). Adult learners also demonstrated a significant negative correlation between Spanish attitude scores and average intra-learner variance ($r(89) = -.450$, $p < .001$). Figure 29 illustrates the positive correlation observed between Spanish attitude scores and average task scores in adult learners. In summary, more positive attitudes about the Spanish language and target language community were significantly correlated with higher accuracy scores and lower average intra-learner variance in adult Spanish learners, and this association was most pronounced on auditory and speeded tasks.

Table 26. Output of a correlation matrix (Pearson's *r*) examining the relationship between attitudes about Spanish and accuracy and intra-learner variance on different task types in **learners** of Spanish. Significant (95% CI) correlations are highlighted in grey.

	SPAN ATTITUDE SCORE	
	r	*p*
AVG TASK SCORE	0.484	< .001
AVG VAR	-0.450	< .001
AVG UNGRAM FEM SCORE	0.335	0.001
SPEEDED SCORE	0.538	< .001
SELF-PACED SCORE	0.169	0.111
AUDITORY SCORE	0.552	< .001
WRITTEN SCORE	0.316	0.002

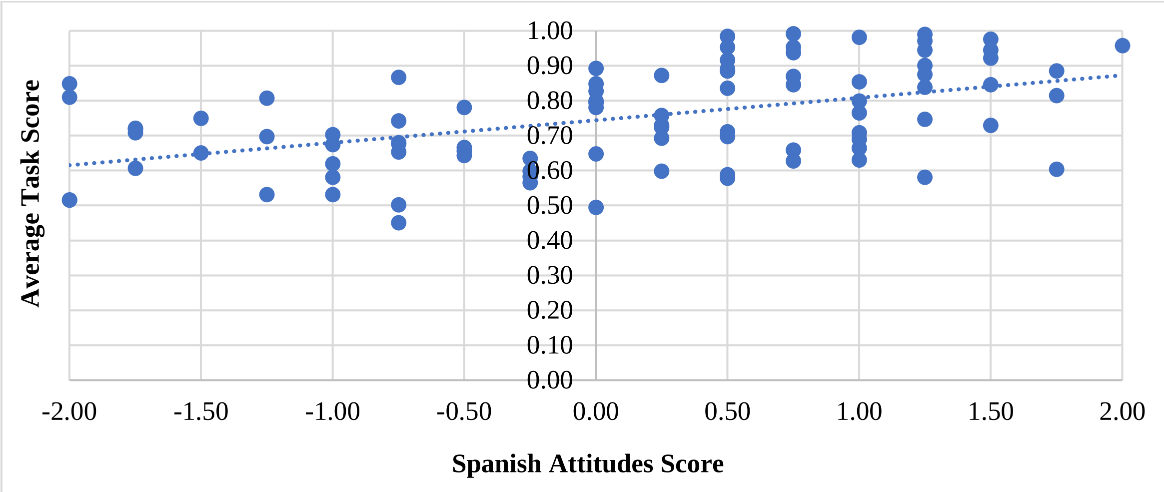
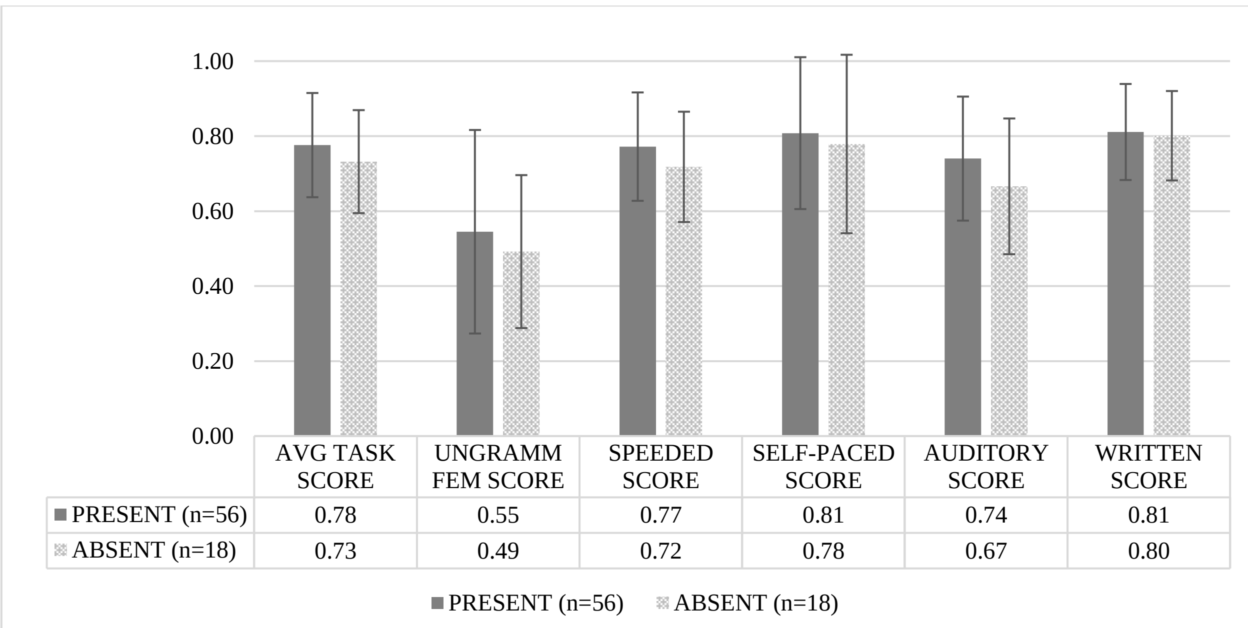


Figure 29. Correlation between average task score and attitudes about Spanish score in **learners**.

5.1.7 Individual factors: prior linguistic repertoire

The effect of prior linguistic repertoire was also investigated. Adult Spanish learners were grouped by the presence or absence of grammatical gender in their linguistic repertoire (other than their current target language of Spanish) according to the other languages they reported in the language history section of the Language Learner Profile Questionnaire. Nineteen different additional languages were reported by participants on the Language Learner Profile Questionnaire and this data is presented in the Methodology section as part of the description of the study participants. Figure 31 presents the observed relationships between prior linguistic repertoire and accuracy scores and variation in adult learners of Spanish. Learners whose prior linguistic repertoire included a gendered language (n = 56) consistently outperformed to varying degrees their peers whose linguistic repertoire did not include another gendered language (n = 18). It is important to note that the vast majority of learners who reported knowing additional languages reported languages that also feature a gender system. Nonetheless, from this unequal sample, we can still observe some key differences between adult learners, particularly on the auditory task in which those learners with grammatical gender *present* in their linguistic repertoire outperformed their peers by a difference of 7% and demonstrated an advantage of 5% on their average speeded task accuracy scores. Furthermore, those learners with grammatical gender *present* performed better (M = .545, SD = .271) than their peers with grammatical gender *absent* in their linguistic repertoire (M = .492, SD = .204) on the more difficult ungrammatical feminine noun tokens. On average task scores, learners with grammatical gender *present* outperformed their peers by an average of 5%. Finally, with regards to group variation (SD), learners with and without grammatical gender present in their prior linguistic repertoire demonstrated comparable variation in their accuracy scores, as can be observed in the error bars featured in Figure 30. In summary, the presence of grammatical gender in learners' prior linguistic repertoire appears to have some effect on their performance with grammatical gender in Spanish as an additional language, and this positive effect is more pronounced on speeded and auditory tasks and appears to afford particular advantage on the more difficult ungrammatical feminine noun tokens. Nonetheless, the advantage that prior linguistic repertoire affords is less evident on group variation.



	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
PRESENT (n=56)	0.78	0.55	0.77	0.81	0.74	0.81
ABSENT (n=18)	0.73	0.49	0.72	0.78	0.67	0.80

Figure 30. **Learners** compared by presence or absence of grammatical gender in prior linguistic repertoire on average task accuracy scores. Standard deviation (SD) bars are provided.

5.1.8 Individual factors combined: predictive modeling

In order to investigate how individual differences together impact performance with grammatical gender in both native Spanish speakers and adult Spanish learners, multiple linear regression models were constructed at a confidence interval of 95%. The final model demonstrating the impact of individual differences in adult learners is provided in Table 27, while Table 28 shows the model for native Spanish speakers.

A separate multiple linear regression model was constructed for just the learner group (n = 90). A significant model was found that accounts for approximately 67% of the variation in accuracy scores among Spanish learners (adjusted $R^2 = 0.672$, $F = 27.1$, $p < .001$), including the individual factors of total Spanish proficiency, metacognitive awareness, metalinguistic awareness of the target structure, overall motivational orientation score, Spanish attitude scores about the target language and target language community, weekly Spanish use, and prior linguistic repertoire (i.e., presence of grammatical gender in additional reported language(s)) of the individual. When the model coefficients are considered individually, the only factors that are significant are total Spanish proficiency ($F = 55.55$, $p < .001$), metalinguistic awareness score ($F = 18.69$, $p < .001$), and motivation ($F = 5.41$, $p = .022$), whereas metacognitive awareness ($F = .66$, $p = .418$), Spanish attitudes ($F = 0.77$, $p = .382$), Spanish use ($F = 0.26$, $p = .614$), and prior linguistic repertoire ($F = .29$, $p = .591$) are non-significant factors, according to the output of an Omnibus ANOVA test, as shown in Table 27. Therefore, in adult Spanish learners, total Spanish proficiency is most predictive of performance with grammatical gender, followed by metalinguistic awareness of the target structure and motivation to learn the language, whereas all other individual factors do not appear to be significantly predictive of performance.

Table 27. Output of a multiple linear regression model (CI 95%) examining the effects of individual differences on average task scores, including Spanish proficiency, metacognitive awareness, metalinguistic awareness, motivation, attitudes about Spanish, weekly Spanish use, and prior linguistic repertoire (presence of grammatical gender in additional reported languages) in **learners**.

				Overall Model Test	
Model	***R***	***R²***	**Adjusted *R²***	***F***	***p***
1	0.835	0.698	0.672	27.1	< .001

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
SPAN PROF: TOTAL SCORE	0.35758	1	0.35758	55.553	< .001
METACOG SCORE	0.00427	1	0.00427	0.664	0.418
METALING AWARE: TOTAL	0.12032	1	0.12032	18.693	< .001
MOTIV: TOTAL SCORE	0.03484	1	0.03484	5.413	0.022
SPAN ATTITUDE SCORE	0.00498	1	0.00498	0.774	0.382

SPAN USE SCORE	0.00165	1	0.00165	0.257	0.614
LING REPERTOIRE	0.00188	1	0.00188	0.291	0.591
Residuals	0.52781	82	0.00644		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept	0.45795	0.0568	0.34487	0.5710	8.056	< .001
SPAN PROF: TOTAL SCORE	0.38863	0.0521	0.28490	0.4924	7.453	< .001
METACOG SCORE	-0.01531	0.0188	-0.05271	0.0221	-0.815	0.418
METALING AWARE: TOTAL	0.13214	0.0306	0.07134	0.1929	4.324	< .001
MOTIV: TOTAL SCORE	0.04116	0.0177	0.00597	0.0764	2.327	0.022
SPAN ATTITUDE SCORE	0.00955	0.0109	-0.01204	0.0311	0.880	0.382
SPAN USE SCORE	-0.05882	0.1161	-0.28979	0.1721	-0.507	0.614
LING REPERTOIRE	0.00995	0.0184	-0.02671	0.0466	0.540	0.591

Native Spanish speakers (n = 25) were considered separately and a multiple linear regression model was constructed to determine to what extent the individual factors investigated in native Spanish speakers may also be predictive of performance. The final model, including total Spanish proficiency, metalinguistic awareness of the target structure, and attitudes about the target language and target language community was *not* significant (adjusted $R^2 = 0.0340$, $F = 1.21$, $p = .337$). Furthermore, none of the individual model coefficients reached significance, according to the output of an Omnibus ANOVA test, as shown in Table 28. Therefore, although some individual differences, namely tested Spanish proficiency, metalinguistic awareness, and motivation do appear to be predictive of performance in adult learners, individual factors are not significantly predictive of performance in native Spanish speakers.

Table 28. Output of a multiple linear regression model (CI 95%) examining the effects of individual differences on average task scores, including Spanish proficiency, metalinguistic awareness, attitudes about Spanish, and weekly Spanish use in **native Spanish speakers** only.

				Overall Model Test	
Model	***R***	***R²***	**Adjusted *R²***	***F***	***p***
1	0.442	0.195	0.0340	1.21	0.337

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
SPAN PROF: TOTAL SCORE	3.61e-4	1	3.61e-4	0.6970	0.414
METALING AWARE: TOTAL	5.56e-4	1	5.56e-4	1.0750	0.312
SPAN ATTITUDE SCORE	6.90e-4	1	6.90e-4	1.3327	0.262
SPAN USE SCORE	1.29e-5	1	1.29e-5	0.0250	0.876
Residuals	0.0103	20	5.17e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept	0.89289	0.05667	0.77468	1.0111	15.756	< .001
SPAN PROF: TOTAL SCORE	0.05497	0.06584	-0.08237	0.1923	0.835	0.414
METALING AWARE: TOTAL	0.02069	0.01995	-0.02093	0.0623	1.037	0.312
SPAN ATTITUDE SCORE	0.01137	0.00985	-0.00917	0.0319	1.154	0.262
SPAN USE SCORE	-0.00369	0.02332	-0.05234	0.0450	-0.158	0.876

5.1.9 Individual factors: inter-factor correlations

Finally, we examine if and to what extent the individual factors analyzed may be correlated with one another, revealing something about the nature of individual differences and how different predictors of learner performance may be intertwined. Table 29 presents the output of a Pearson's *r* correlation matrix between the different individual factors considered among adult learners. The strongest correlations among individual learner factors were observed between Spanish attitudes and motivation ($r(89) = .554$, $p < .001$), between Spanish proficiency and Spanish attitudes ($r(89) = .518$, $p < .001$), between metacognitive awareness and motivation ($r(89) = .410$, $p < .001$), and between Spanish proficiency and motivation ($r(89) = .408$, $p < .001$). Spanish use was also found to be moderately correlated with proficiency, attitudes, and motivation as was metalinguistic awareness with proficiency and motivation. Interestingly, regarding learner awareness, metacognitive awareness and metalinguistic awareness were *not* found to be significantly correlated. Furthermore, metacognitive awareness was *not* found to be significantly correlated with Spanish proficiency. Nonetheless, many of the analyzed individual factors are clearly intertwined and may develop together in unison, affecting learner performance in the language at any point in time.

Table 29. Output of a correlation matrix (Pearson's *r*) examining the relationship between different individual factors in **Spanish learners**. The individual factors analyzed include: average tested Spanish proficiency, attitudes about the Spanish language and target language community, motivation level, use of Spanish across different contexts in an average week, metacognitive awareness about the language learning process and associated strategies, metalinguistic awareness of grammatical gender, and prior linguistic repertoire (presence of grammatical gender in additional reported languages). Significant (95% CI) correlations are highlighted in grey.

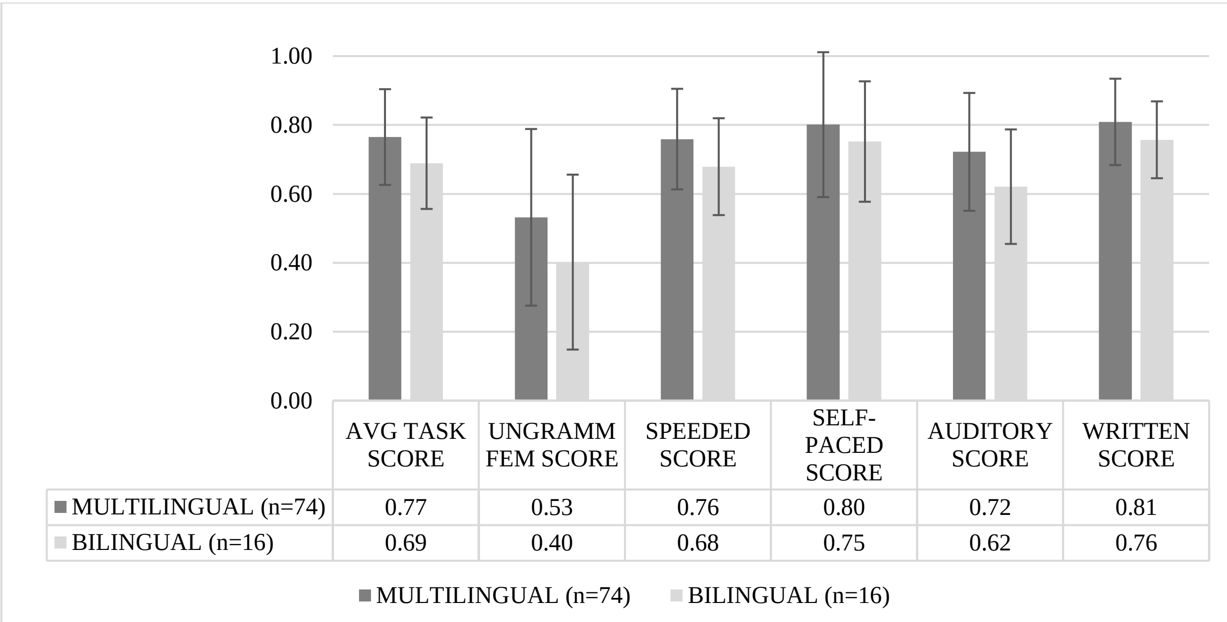
		SPANISH PROFICIENCY	SPANISH ATTITUDES	MOTIVATION	SPANISH USE	METACOG AWARENESS	METALING AWARENESS
SPANISH PROFICIENCY	*r*	—					
	p	—					
SPANISH ATTITUDES	*r*	0.518	—				
	p	< .001	—				
MOTIVATION	*r*	0.408	0.554	—			
	p	< .001	< .001	—			
SPANISH USE	*r*	0.344	0.346	0.366	—		
	p	< .001	< .001	< .001	—		
METACOG AWARENESS	*r*	0.143	0.278	0.410	0.034	—	
	p	0.178	0.008	< .001	0.752	—	
METALING AWARENESS	*r*	0.292	0.118	0.216	0.107	0.163	—
	p	0.005	0.269	0.040	0.314	0.126	—
LING REPERTOIRE	*r*	0.232	0.028	0.175	-0.004	0.135	0.084
	p	0.028	0.793	0.098	0.972	0.206	0.431

5.2 Multilingual effect

The fourth and final research question examined how being multilingual when learning Spanish as an additional language may affect performance with grammatical gender in adult learners. Furthermore, the extent to which knowing another gendered language may explain any effect of multilingualism was also explored as well as the effect of the number of additional languages reported by multilingual learner participants. As a point of comparison, the effect of multilingualism was also explored in the native speaker group. To examine the effect of multilingualism on performance with grammatical gender in Spanish, all participants (N = 115) were grouped according to whether or not they reported knowledge of any additional language(s) beyond English and Spanish, forming two sub-groups: “multilinguals” (n = 86) and “bilinguals” (n = 29). A full description of the presence of multilingualism in the present sample is presented in the Methodology section. For further analysis, learners (n=90) and native speakers (n=25) were subsequently grouped according to their self-reported multilingual status. In what follows, the descriptive results will be presented in detail for the effect of multilingualism and then the results of inferential analysis using multiple linear regression modeling will be presented to substantiate the descriptive observations.

5.2.1 Multilingual effect: general advantage

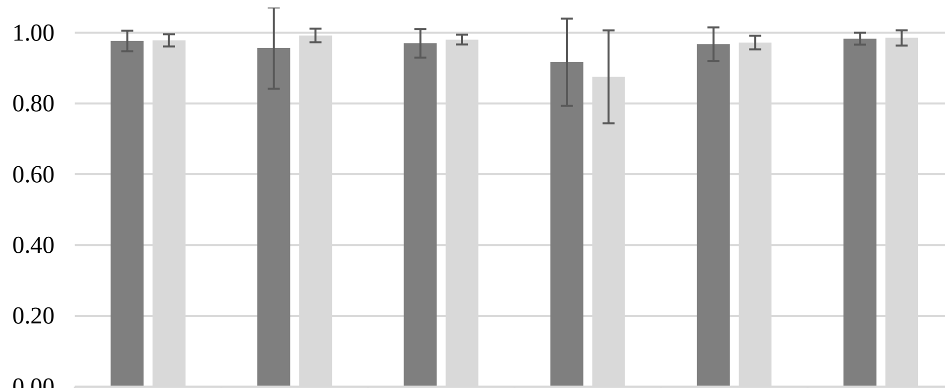
The learner group was subdivided according to whether or not they reported knowledge of any additional language(s) beyond English and Spanish, forming two groups: “multilingual learners” (n = 74) and “bilingual learners” (n = 16). The descriptive results of this analysis are presented in Figure 31. Overall, it appears that being multilingual as an adult learner of Spanish affords certain advantage as multilingual learners consistently outperformed their bilingual peers on all measures, producing an 8% advantage in average task scores for multilingual learner participants. The multilingual learner advantage was most pronounced on average scores on ungrammatical feminine nouns in which multilingual learners (*M* = .532, *SD* = .256) outperformed their bilingual learner peers (*M* = .402, *SD* = .254) by 13%. In addition, multilingual learners scored notably higher on auditory tasks (*M* = .722, *SD* = .171) than bilingual learners (*M* = .621, *SD* = .166), and on average scored higher on speeded tasks (*M* = .759, *SD* = .146) than their bilingual learner peers (*M* = .679, *SD* = .141). In terms of group variation (SD) in performance, multilingual and bilingual learners as a group were very similar, with the exception of average self-paced task scores in which multilingual learners actually demonstrated greater variation in their scores than their bilingual peers, as observed in the error bars shown in Figure 31. In summary, being multilingual affords certain advantage for learners as it is associated with higher accuracy rates, particularly with the more difficult ungrammatical feminine noun token type, and higher scores on the more cognitively demanding auditory and speeded tasks.



	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
MULTILINGUAL (n=74)	0.77	0.53	0.76	0.80	0.72	0.81
BILINGUAL (n=16)	0.69	0.40	0.68	0.75	0.62	0.76

Figure 31. Multilingual vs bilingual groups compared on average task accuracy scores in **learners** only. Standard deviation (SD) bars are provided.

As a point of comparison, the effect of multilingualism was investigated separately in native Spanish speakers who were subdivided into two nearly equal groups: “multilingual native speakers” (n = 12) and “bilingual native speakers” (n = 13). The descriptive results of this analysis are presented in Figure 32. Overall, *bilingual* native speakers demonstrated a slight advantage over their multilingual peers on half of the measures analyzed and this slight bilingual advantage was most pronounced on average ungrammatical feminine noun scores in which bilingual native speakers (M = .992, SD = .019) scored 3% higher on average than their multilingual native speaker peers (M = .957, SD = .115). However, with regards to average self-paced task scores, multilingual native speakers (M = .917, SD = .123) performed better than their bilingual counterparts (M = .875, SD = .131). Nonetheless, the differences in average accuracy scores between multilingual and bilingual native speaker participants were quite minimal on most measures. With regards to group variation (SD) in performance, multilingual native speakers actually demonstrated *more* variation on most measures than their bilingual peers, as shown in the error bars in Figure 32. In summary, the effect of multilingualism was rather minimal among native speakers and only afforded a slight advantage on self-paced task scores yet was associated with an even slighter disadvantage on other accuracy scores and is further associated with greater variation in performance.



	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
MULTILINGUAL (n=12)	0.98	0.96	0.97	0.92	0.97	0.98
BILINGUAL (n=13)	0.98	0.99	0.98	0.88	0.97	0.99

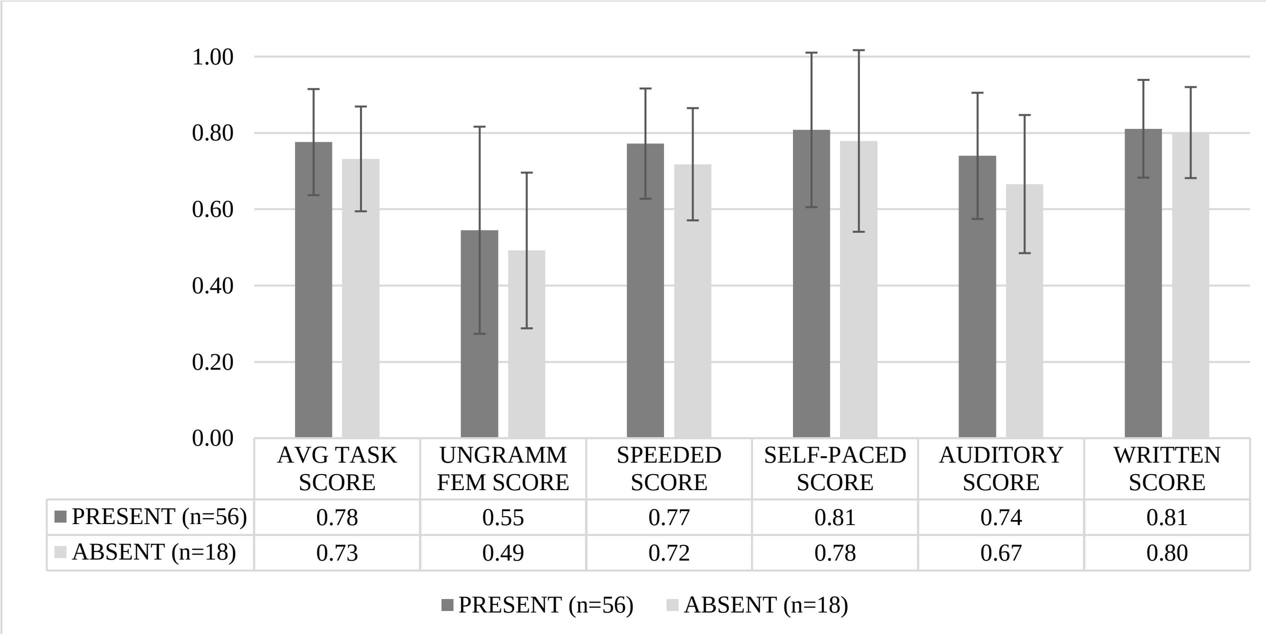
MULTILINGUAL (n=12) BILINGUAL (n=13)

Figure 32. Multilingual vs bilingual groups compared on average task accuracy scores in **native Spanish speakers** only. Standard deviation (SD) bars are provided.

5.2.2 Multilingual effect: general typology

In order to investigate more precisely the nature of the positive effect of multilingualism in adult learners of Spanish, multilingual learners (N = 74) were further subdivided according to whether or not they reported knowledge of an additional language (beyond Spanish) that also exhibits a grammatical gender system, producing two subgroups: multilingual learners with grammatical gender “present” in their prior linguistic repertoire (n = 56) and multilingual learners with grammatical gender “absent” in their prior linguistic repertoire (n = 18). This subdivision of multilingual learner participants allowed us to explore the question of whether or not knowing another gendered language may be independent from the effect of being multilingual. In other words, this analysis allows us to consider to what extent the observed multilingual advantage in adult learners may be attributable to the transfer of a pre-existing grammatical gender system in the prior linguistic repertoire of the individual or to what extent the multilingual effect may be independent of language typology with regards to gender. The descriptive results of this analysis are presented in Figure 33. Although approximately half (12/25) of native Spanish speakers reported being multilingual, native speakers were excluded from this analysis due to the fact that their native language features a grammatical gender system and there is no objective way to tease apart the effect of a gendered L1 from the effect of a gendered L*n* (other additional language) in the present sample.

Overall, the effect of a grammatical gender system in the prior linguistic repertoire of multilingual learner participants is consistently positive on all measures analyzed; those multilingual learners with grammatical gender "present" in their prior linguistic repertoire outperformed their peers by 5% on average task scores, and this advantage was most pronounced on auditory scores, producing a 7% advantage for multilingual learners with grammatical gender "present" (M = .740, SD = .165) as compared to multilingual learners with grammatical gender "absent" (M = .666, SD = .181) in their prior linguistic repertoire. Nearly equal advantage was also observed for multilingual learners with grammatical gender "present" for their average speeded task scores (M = .772, SD = .145) and their average scores on ungrammatical feminine nouns (M = .545, SD = .271), compared to their peers with grammatical gender "absent" from their prior linguistic repertoire (speeded score: M = .718, SD = .147; ungrammatical feminine noun score: M = .492, SD = .204). In addition, multilingual learners with grammatical gender "present" demonstrated slightly higher accuracy on self-paced tasks (M = .808, SD = .203) than their peers (M = .779, SD = .238). However, with regards to written scores as well as with regards to their respective group variation (SD) in performance, all multilingual participants performed in a remarkably similar way on tasks, regardless of the status of grammatical gender in their prior linguistic repertoire. In summary, the positive effect of multilingualism in adult learners appears to be connected to the typology of their prior linguistic repertoire in that multilingual learners with grammatical gender "present" in their repertoire exhibit higher accuracy rates, particularly on auditory and speeded tasks, and higher accuracy on the more difficult ungrammatical feminine noun token type than their multilingual learner peers whose prior linguistic repertoire does not feature a grammatical gender system. Figure 33 illustrates the typological multilingual learner advantage across the different measures analyzed.

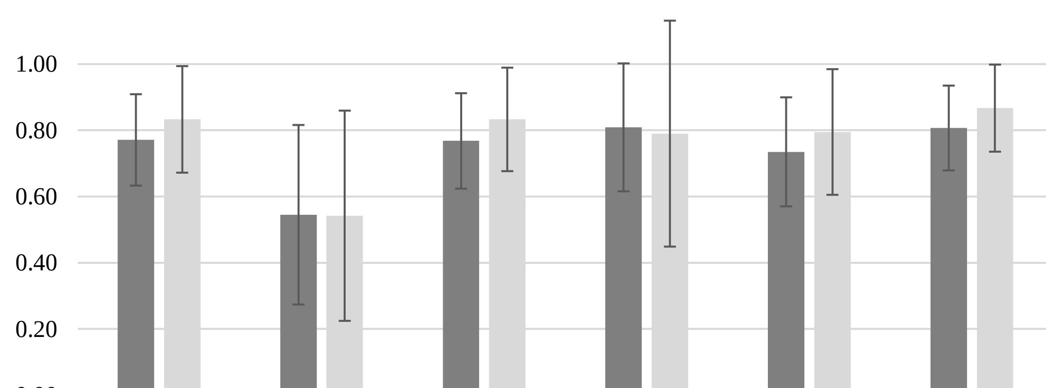


	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
PRESENT (n=56)	0.78	0.55	0.77	0.81	0.74	0.81
ABSENT (n=18)	0.73	0.49	0.72	0.78	0.67	0.80

Figure 33. Multilingual **learner** participants compared by **presence of grammatical gender** in their reported prior linguistic repertoire. Standard deviation (SD) bars are provided.

5.2.3 Multilingual effect: gendered language subtype

Given the observation in adult learners of Spanish that being multilingual is associated with an advantage in performance, and in addition, given that the presence of a gendered language in the prior linguistic repertoire of the multilingual learner affords additional advantage, it is also relevant, then, to examine the effect of gendered language subtype, that is, the presence of *binary* grammatical gender in the linguistic repertoire of the multilingual learner as compared to other manifestations of grammatical gender into more than two noun class categories. For this analysis, multilingual learners whose prior linguistic repertoire includes another gendered language (N = 56) were further subdivided according to the gender *type* of their additional language(s) reported, forming two groups: multilingual learners with *binary* grammatical gender "present" in their prior linguistic repertoire (n = 52), and multilingual learners with *binary* grammatical gender "absent" in their prior linguistic repertoire (n = 4). The descriptive results of this analysis are presented in Figure 34. Overall, the presence of *binary* grammatical gender in the multilingual learner's prior linguistic repertoire does *not* appear to afford any additional advantage in performance with grammatical gender in Spanish as an additional language; in fact, multilingual learners with binary grammatical gender "absent" in their linguistic repertoire outperformed their peers by 6% on average task scores. Furthermore, multilingual learners with binary grammatical gender "absent" (*M* = .833, *SD* = .156) outperformed their peers (*M* = .768, *SD* = .144) by 6% on speeded tasks as well. In addition, the *negative* effect of the presence of binary grammatical gender in the prior linguistic repertoire appears to be independent of task stimuli modality as those with binary grammatical gender "absent" equally outperformed their peers on both written and auditory tasks by an average of 6%. The only measure on which a slight advantage was observed for those multilingual learners with binary grammatical gender "present" was average self-paced scores in which they outperformed (*M* = .809, *SD* = .193) their peers with grammatical gender "absent" (*M* = .790, *SD* = .341) by 2%. Furthermore, the presence of binary grammatical gender did not appear to have a differentiating effect with average scores on ungrammatical feminine nouns, in which multilingual learners with binary grammatical gender "present" (*M* = .545, *SD* = .271) had nearly identical scores to their peers (*M* = .542, *SD* = .318). Finally, with regards to group variation (SD) in performance, multilingual learners both with and without binary grammatical gender present were very similar, except for average self-paced scores in which multilingual learners with binary grammatical gender *absent* demonstrated notably higher variation in their performance, as shown in the error bars in Figure 34.



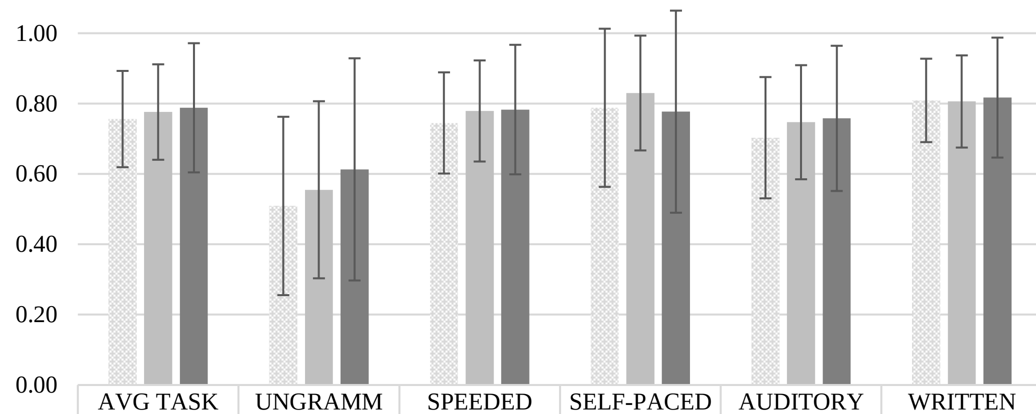
	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
PRESENT (n=52)	0.77	0.55	0.77	0.81	0.74	0.81
ABSENT (n=4)	0.83	0.54	0.83	0.79	0.80	0.87

Figure 34. Multilingual **learner** participants compared by presence of **binary grammatical gender** in their reported prior linguistic repertoire. Standard deviation (SD) bars are provided.

In summary, the typological multilingual learner advantage appears to be limited to the presence of a grammatical gender system in the prior linguistic repertoire of the multilingual learner, but does not, however, appear to be associated with the particular type of gendered noun class system; multilingual learners with *binary* grammatical gender present in their prior linguistic repertoire did *not* perform better than their peers with other types of gendered noun class systems, and, in fact, the presence of binary grammatical gender appears to be associated with lower average accuracy rates, particularly on speeded, auditory and written tasks and does not appear to reduce group variation in performance. Finally, no effect of noun gender class system type is observed on the more difficult ungrammatical feminine noun tokens.

5.2.4 Multilingual effect: number of additional languages

Given the observed positive effect that multilingualism has on adult learners' performance with grammatical gender in Spanish as an additional language, it is also pertinent to examine if the number of additional languages known may differentiate this effect. In other words, we examine here if knowing two or three additional languages affords a greater advantage than knowing just one additional language. For this analysis, multilingual learner participants (N = 74) were subdivided into three groups, according to the number of additional languages reported beyond English and Spanish: "*one (1)* additional language" (n = 44), "*two (2)* additional languages" (n = 24), and "*three (3)* additional languages" (n = 6). The descriptive results of this analysis are presented in Figure 35. Overall, speaking three (3) additional languages afforded consistent advantage over speaking just one or two additional languages on almost all measures analyzed. This advantage was most pronounced on average ungrammatical feminine noun scores in which participants reporting three additional languages outperformed (*M* = .613, *SD* = .316) their peers reporting only two (*M* = .555, *SD* = .252) or one (*M* = .509, *SD* = .254) additional language(s). A notable advantage of knowing three additional languages was also observed on auditory task scores, resulting in a 6% advantage for multilingual learners reporting three additional languages (*M* = .758, *SD* = .206) compared to their peers reporting two (*M* = .747, *SD* = .162) and one (*M* = .703, *SD* = .172) additional language(s). Although multilingual learners who know three additional languages outperformed their other multilingual peers by 3% on average task scores, this advantage was less clear on self-paced and written tasks; on written tasks, average scores were remarkably similar for all three groups, and on the self-paced tasks, accuracy scores were actually higher for those reporting two additional languages (M = .830, SD = .163) than those multilingual learners who reported three additional languages (M = .777, SD = .287).



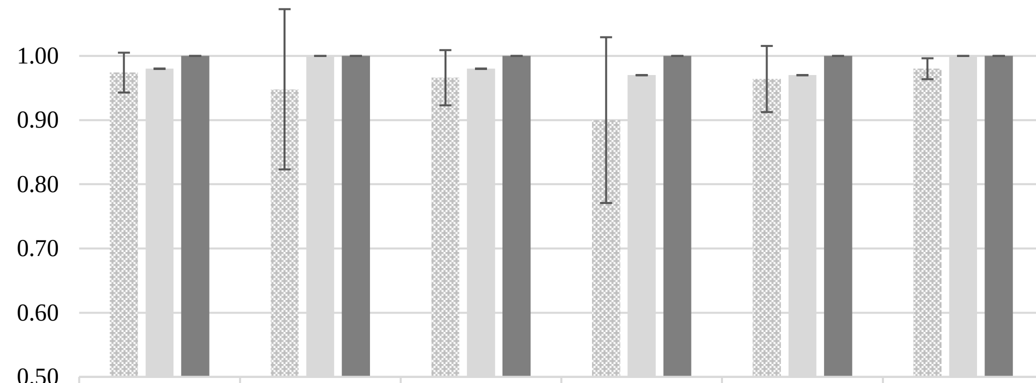
	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
1 ADD LANG (n=44)	0.76	0.51	0.75	0.79	0.70	0.81
2 ADD LANG (n=24)	0.78	0.56	0.78	0.83	0.75	0.81
3 ADD LANG (n=6)	0.79	0.61	0.78	0.78	0.76	0.82

1 ADD LANG (n=44) 2 ADD LANG (n=24) 3 ADD LANG (n=6)

Figure 35. Multilingual **learner** participants compared by number of additional languages reported. Standard deviation (SD) bars are provided.

Furthermore, the difference in accuracy scores between knowing three additional languages and two additional languages was quite minimal across nearly all measures analyzed. In addition, group variation (SD) in performance remained rather stable across learner groups. In summary, the number of additional languages known appears to have an effect on multilingual adult learner performance in which knowing two or more additional languages is more advantageous than knowing just one additional language, and this effect is particularly pronounced with the more difficult ungrammatical feminine noun token type and with the more cognitively demanding auditory tasks. Nonetheless, the effect of the number of additional languages known on self-paced and written tasks is less evident and appears to be more prominent when comparing multilinguals reporting one versus three additional languages, whereas much less of an effect is observed between multilingual participants reporting two and three additional languages. Therefore, a more robust advantage is observed for knowing *two or more* additional languages as compared to just one additional language in multilingual learners of Spanish.

The effect of the number of additional languages reported was also investigated in the multilingual native speaker group (N = 12) who were subdivided, just like the multilingual learner group, into three subgroups according to the number of additional languages they reported beyond Spanish and English: "*one (1)* additional language" (n = 10), "*two (2)* additional languages" (n = 1), and "*three (3)* additional languages" (n = 1). The present analysis is particularly limited by the fact that these subgroups are very unequal in that most multilingual native speakers reported knowledge of only one additional language. Nonetheless, the descriptive results of this analysis are presented in Figure 36, for which further data is needed in order to confirm the observed trends. Despite the obvious limitations of the sample, the multilingual native speaker (n=1) who reported knowledge of three additional languages appears to have outperformed their peers who reported only one additional language on every measure analyzed, although this may be due to chance. This apparent advantage was most pronounced on average self-paced task scores resulting in a 10% difference compared to their peers reporting one additional language. In addition, average scores on ungrammatical feminine nouns resulted in the second largest difference between the multilingual native speaker reporting three additional languages (*M* = 1.00, *SD* = 0.0) and those speakers reporting only one additional language (*M* = .948, *SD* = .125). Nonetheless, the other native speaker who reported two additional languages also performed at ceiling (*M* = 1.00) on their average self-paced task score, average written task score, and average ungrammatical feminine noun score. Standard deviation (SD) as a measure of variation in group performance could only be calculated for the multilingual native speaker subgroup reporting one additional language (n=10), since the other two groups only contained one individual and therefore standard deviation cannot be calculated. For the one additional language group, the most variation was observed on ungrammatical feminine noun scores and on average self-paced task scores, as shown in the standard deviation bars represented in Figure 36.



	AVG TASK SCORE	UNGRAMM FEM SCORE	SPEEDED SCORE	SELF-PACED SCORE	AUDITORY SCORE	WRITTEN SCORE
1 ADD LANG (n=10)	0.97	0.95	0.97	0.90	0.96	0.98
2 ADD LANG (n=1)	0.98	1.00	0.98	0.97	0.97	1.00
3 ADD LANG (n=1)	1.00	1.00	1.00	1.00	1.00	1.00

1 ADD LANG (n=10)　2 ADD LANG (n=1)　3 ADD LANG (n=1)

Figure 36. Multilingual **native speaker** participants compared by number of additional languages reported. Standard deviation (SD) bars are provided. *Note: due to a sample size of one (1) for the categories of “2 ADD LANG” and “3 ADD LANG”, the standard deviation of the group is zero (‘0’).

In summary, this limited sample of multilingual native speakers seems to show that knowing three additional languages as compared to just one additional language has a positive effect as it is associated with higher accuracy scores, particularly on self-paced tasks and on ungrammatical feminine noun tokens. However, the effect of ‘number of additional languages’ appears to pertain more to ‘one additional language’ versus ‘more than one’, similar to the binary trend observed in the multilingual learner group (See Figure 35). Nonetheless, it must be noted that the present analysis of the effect of number of additional languages reported in native speaker performance is extremely limited and inconclusive due to the uneven groupings identified in the present sample of native Spanish speakers. Therefore, further research is needed to confirm these preliminary descriptive findings.

5.2.5 Multilingual effect: predictive modeling

In order to determine to what extent the effect of multilingualism observed in the present sample of both native speakers and adult learners may be significant, linear regression models were constructed at a confidence interval of 95%. The final model demonstrating the effect of multilingualism on average task scores in adult learners is presented in Table 30 and the final regression model for native Spanish speakers is presented in Table 31.

A significant linear regression model was found for the learner group (n = 90) in which the status of grammatical gender ("present" – "absent") in the individual's prior linguistic repertoire accounts for approximately 4% of the variation in accuracy scores among Spanish learners (adjusted $R^2 = 0.038$, $F = 4.51$, $p = .036$). This was the most significant model compared to models that also included the coefficients 'multilingualism' and 'binary grammatical gender'. The output of an Omnibus ANOVA test with 'grammatical gender present' as the only model coefficient is presented in Table 30. In sum, the significant linear regression model found indicates that accuracy scores with grammatical gender in Spanish can be partially accounted for—to a very minimal extent (4%)—by the presence of a grammatical gender system in the prior linguistic repertoire of the multilingual learner. Evidently, the other individual factors discussed previously contribute notably more to learner performance.

Table 30. Output of a linear regression model (CI 95%) examining the effect of multilingualism on average task scores for **learners** only.

				Overall Model Test	
Model	**R**	**R²**	**Adjusted R²**	***F***	***p***
1	0.221	0.0488	0.0380	4.51	.036

Omnibus ANOVA Test

	Sum of Squares	df	Mean Square	*F*	*p*
GRAMMATICAL GENDER PRESENT:	0.0852	1	0.0852	4.51	.036
Residuals	1.6623	88	0.0189		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	0.7755	0.0184	0.739	0.812	42.23	< .001

GRAMMATICAL GENDER PRESENT:						
absent – present	-0.0635	0.0299	-0.123	-0.00409	-2.12	0.036

Lastly, the predictive power of the 'multilingualism' factor was investigated in native Spanish speakers (n = 25) through linear regression modeling. However, the final linear regression model with 'multilingualism' as the only model coefficient to predict average task score did not reach significance (adjusted R^2 = -0.042, F = .036, p = .851), as presented in Table 31. Therefore, it appears that in native Spanish speakers, the effect of multilingualism is not significant and is not predictive of performance with grammatical gender in Spanish as a native language.

Table 31. Output of a linear regression model (CI 95%) examining the effect of multilingualism on average task scores for **native Spanish speakers** only.

				Overall Model Test	
Model	**R**	**R²**	**Adjusted R²**	***F***	***p***
1	0.0395	0.00156	-0.0418	0.0360	.851

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
MULTILINGUALISM	2.01e-5	1	2.01e-5	0.0360	0.851
Residuals	0.0128	23	5.58e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	0.97667	0.00682	0.9626	0.9908	143.215	< .001
MULTILINGUALISM:						
No – Yes	0.00179	0.00946	-0.0178	0.0214	0.190	0.851

5.3 Summary of results: Individual factors and the multilingual effect

With regards to the impact of individual factors, multiple linear regression modeling revealed that total Spanish proficiency is most predictive of adult learner performance with grammatical gender, followed by metalinguistic awareness of the target structure and motivation to learn the language, whereas all other individual factors were not found to be significantly and independently predictive of adult learner performance, although significant positive correlations were found between accuracy scores and metacognitive awareness, language use, attitudes, and linguistic repertoire in adult learners. In contrast, the individual factors of proficiency, metalinguistic awareness, attitudes, and language use were not found to be significantly predictive of performance in native Spanish speakers. Furthermore, many of the analyzed individual factors were found to be significantly correlated and therefore may develop together in unison, affecting learner performance in the target language at any point in time. More specifically, inter-factor correlations indicated that more motivated learners tend to have more positive attitudes about the target language community and also tend to be more aware of their own language learning process and associated strategies. Moreover, more proficient learners tend to have more positive attitudes toward the target language community and demonstrate higher levels of motivation to learn the language. Inter-factor correlations also revealed that as target language proficiency increases, use of the language, positive attitudes about the language, and motivation to learn the language also increase, and explicit knowledge about the language (i.e., metalinguistic awareness) is also enhanced in adult learners as their proficiency develops.

A slight typological and quantitative multilingual learner advantage was detected in the present findings. Descriptive results demonstrated a consistent multilingual advantage for adult learners across all measures analyzed, and this advantage appeared to be linked to the presence of a grammatical gender system in the prior linguistic repertoire of the adult learner and also appeared to be somewhat linked to multilingual learners reporting two or more additional languages as opposed to just one additional language beyond English and Spanish. Nonetheless, multiple linear regression modeling revealed that accuracy scores with grammatical gender in Spanish as a late acquired nonnative language can only be minimally accounted for (4%) by the presence of a grammatical gender system in the prior linguistic repertoire of the multilingual adult learner. However, the other individual factors of proficiency, metalinguistic awareness, and motivation contribute notably more to learner performance than having a multilingual language background. Although results indicated that multilingualism is partially predictive of learner performance, multilingualism is *not* significantly predictive of performance with grammatical gender in Spanish as a native language.

In the following chapter, the key results will be summarized and discussed in light of the initial predictions made and possible explanations will be considered within the larger context of previous theory and research in adult second language acquisition and multilingualism. Limitations of the present study will also be discussed, and avenues of future research will be considered based on the findings and limitations of this study.

Chapter 6: Discussion

6.1 General Discussion

Recall that the aim of this dissertation study is to examine how linguistic variables, task demands, and individual factors, including a multilingual background and prior linguistic repertoire, all variably modulate, account for, and/or predict performance with grammatical gender—measured in mean accuracy and intra-speaker variance—in both native Spanish speakers and adult learners of Spanish. Through this orientation, we endeavor to identify and better understand what factors are most important in nonnative language acquisition and consider to what extent these same factors may also affect native speaker performance. In addition, we seek to draw principled conclusions regarding to what extent learners and native speakers access different forms of linguistic knowledge—implicit/procedural and explicit/declarative. In order to address these aims, 115 language users of Spanish from diverse language backgrounds were recruited, including 25 native Spanish speakers and 90 late/adult Spanish learners, who all completed a Language Learner Profile questionnaire, a Spanish proficiency test, four experimental tasks in which the task conditions were manipulated (1. speeded oral imitation task; 2. speeded auditory grammaticality judgment task; 3. speeded written grammaticality judgment task; and 4. self-paced written grammaticality judgment task), and a metalinguistic awareness exit survey.

Overall, the results of this study indicate that adult/late learners are differentially affected—in terms of both accuracy and variation—by many of the explanatory variables examined including linguistic, task, and individual factors, whereas native speaker accuracy scores were only significantly modulated by task effects and noun frequency. Nonetheless, both individual intra-learner/speaker variance and group variation (SD) were found to be differentially modulated by these factors in both adult learners and native speakers alike, indicating how late learner and native speaker linguistic performance (both accuracy and variation) can be variably modulated by external factors, contradicting the prototypical and invariable native speaker norm. Furthermore, significant interactions were detected with learner proficiency indicating that factors modulating nonnative performance diminish with increasing proficiency and late learners at the advanced proficiency level can mirror native speaker norms of performance.

With regards to the effect of the linguistic variables, although native speaker accuracy scores were largely unaffected by the linguistic factors, clear distinctions were observed in accuracy scores in adult learners who exhibited significantly enhanced accuracy with frequent, grammatical, and masculine target tokens. Descriptive results indicated that native speakers were slightly affected by noun frequency while learners were much more affected. Proficiency level in adult learners was also found to interact with the relative effect of the linguistic variables analyzed as the advanced proficiency learner group demonstrated *qualitatively* the same pattern in their performance as the lower proficiency groups—showing more accurate performance on grammatical, masculine, and overtly marked tokens—yet *quantitatively* the advanced learner scores were more similar to the native speaker group than to the intermediate learner group as advanced learners approximated native speaker norms of performance. When considered all together through multiple linear regression modeling, relative noun frequency was found to be the most predictive factor of learner performance with grammatical gender, followed by token grammaticality and noun gender class, whereas noun morphology did not appear to be independently predictive of performance in adult learners. In contrast, none of the linguistic factors were found to be individually predictive of performance with grammatical gender in Spanish as a native language.

Results also pointed to task effects, indicating that learners were clearly affected by task type and performed better on the self-paced tasks and on the written tasks, although relative task effects varied per learner proficiency level such that advanced learners patterned very closely—both *quantitatively* and *qualitatively*—with their native speaker counterparts. In contrast to the intermediate and beginner learner groups, both the native speakers and the advanced learners actually performed better on the speeded tasks and showed virtually no difference in their accuracy scores between written and auditory modality tasks. According to multiple linear regression modeling, the time constraint factor was the most powerful predictor of performance in adult learners, followed by the stimuli modality task factor, and when both task factors were considered together, they were highly predictive of performance with grammatical gender in Spanish as a nonnative language. The time constraint task factor was also predictive of performance in native speakers, although the nature of the time constraint effect was reversed for both the advanced learners and the native speakers as they performed significantly better under a time constraint, contrary to their intermediate and beginner learner peers. Furthermore, task stimuli modality was *not* independently predictive of performance with grammatical gender in Spanish as a native language nor at advanced levels of learner proficiency, contrary to the trend observed in lower proficiency learners.

With regards to the impact of individual factors, multiple linear regression modeling revealed that total Spanish proficiency is most predictive of adult learner performance with grammatical gender, followed by metalinguistic awareness of the target structure and motivation to learn the language, whereas all other individual factors were not found to be significantly and independently predictive of adult learner performance, although significant positive correlations were found between accuracy scores and metacognitive awareness, language use, attitudes, and linguistic repertoire in adult learners. In contrast, the individual factors of proficiency, metalinguistic awareness, attitudes, and language use were not found to be significantly predictive of performance in native Spanish speakers. Furthermore, many of the analyzed individual factors were found to be significantly correlated and therefore may develop together in unison, affecting learner performance in the target language at any point in time. More specifically, inter-factor correlations indicated that more motivated learners tend to have more positive attitudes about the target language community and also tend to be more aware of their own language learning process (i.e., metacognitive awareness). Moreover, more proficient learners tend to have more positive attitudes toward the target language community and demonstrate higher levels of motivation to learn the language. Inter-factor correlations also revealed that as target language proficiency increases, use of the language, positive attitudes about the language, and motivation to learn the language also increase, and explicit knowledge about the language (i.e., metalinguistic awareness) is also enhanced in adult learners as their proficiency develops.

A slight typological and quantitative multilingual learner advantage was detected in the present findings. Descriptive results demonstrated a consistent multilingual advantage for adult learners across all measures analyzed, and this advantage appeared to be linked to the presence of a grammatical gender system in the prior linguistic repertoire of the adult learner and appeared to be somewhat linked to multilingual learners reporting two or more additional languages as opposed to just one additional language beyond English and Spanish. Nonetheless, multiple linear regression modeling revealed that accuracy scores with grammatical gender in Spanish as a late acquired nonnative language can only be minimally accounted for (4%) by the presence of a grammatical gender system in the prior linguistic repertoire of the multilingual adult learner. However, the other individual factors of proficiency, metalinguistic awareness, and motivation contribute notably more to learner performance than having a multilingual language background. Although results indicated that multilingualism is partially predictive of learner performance, multilingualism is *not* significantly predictive of performance with grammatical gender in Spanish as a native language.

We will now examine more closely the results as they respond to each research question, determine to what extent these results may corroborate or deviate from our initial predictions, and consider possible explanations for these findings grounded in previous research and theory.

6.2 Native and nonnative performance with gender agreement: Effect of linguistic variables (RQ1)

The first research question asked how native Spanish speakers and late learners of Spanish perform with grammatical gender and considered how the inherent linguistic variables (noun gender class, morphological marking, domain of agreement, and target noun frequency) may modulate performance in these two language user groups. It was predicted that native speakers would perform at or near ceiling and that adult learners would perform above chance level, yet significantly below their native speaker peers. Furthermore, it was predicted that the linguistic variables would impact the accuracy rates of both native speakers and adult learners, but that learners would be significantly *more* affected and would demonstrate higher accuracy scores with overtly marked masculine nouns. Regarding interactions with Spanish proficiency, it was also predicted that as proficiency increases, the relative effect of the linguistic variables would decrease. With regards to variation, it was predicted that adult learners would show significantly higher *inter*-speaker (expressed as the standard deviation of the group) and *intra*-speaker (expressed as the variance detected in each participant's responses) variation. These predictions were largely corroborated by the results.

The average native speaker accuracy rate on all experimental tasks combined was 98% while the average learner accuracy rate was 75%. Regarding individual variation in performance, measured as intra-learner/speaker variance, native speakers consistently demonstrated very low variance in their scores, with minor fluctuations on Task 3 (speeded auditory grammaticality judgment task) and Task 6 (metalinguistic awareness exit survey), while learners demonstrated consistently higher variation across all tasks, both as a group (SD) and individually (intra-learner variance).

Regarding the relative effect of each of the linguistic variables investigated, while native speaker accuracy scores remained relatively consistent across the linguistic variable measures, clear distinctions were observed in accuracy scores in adult learners who exhibited enhanced accuracy in their performance with grammatical, masculine, and overtly marked target tokens. Furthermore, the linguistic variables analyzed produced a much larger effect size on accuracy scores in the learner group, whereas native speakers were largely unaffected by these variables. More precisely, the two main participant groups were most distinct from one another on their accuracy with ungrammatical, feminine, non-overtly marked tokens with an average learner group accuracy score of 41% whereas their native speaker counterparts scored a 96% on these same tokens when averaged across all tasks.

The effect of target noun frequency was also analyzed and compared for each participant group. Native speakers' performance was only slightly affected by noun frequency whereas learners were much more affected, demonstrating an average difference in accuracy scores of nearly 5% between high and low frequency target nouns, with a correspondingly higher Pearson's *r* correlation coefficient of .06 for learners versus only .02 for native speakers. Nonetheless, the correlations between accuracy and noun frequency are notably weak in both groups.

In order to examine to what extent proficiency in Spanish may influence the relative effect on performance of the linguistic variables analyzed, accuracy scores were compared across learner proficiency groups and the native speaker group. Results indicate that although the advanced proficiency learner group demonstrated *qualitatively* the same pattern in their performance as the lower proficiency groups, including intermediate and beginner, showing greater accuracy with grammatical, masculine, and overtly marked tokens, *quantitatively* the advanced learner scores were more similar to the native speaker group than to the intermediate learner group, thereby demonstrating that as proficiency increases, the effect that linguistic variables have on learner performance diminishes and advanced proficiency learners can very closely approximate native speaker norms of performance.

To determine to what extent the descriptive findings could be corroborated by inferential analysis, multiple linear regression models with a confidence interval of 95% were also constructed. For adult Spanish learners, a significant multiple linear regression model was found that accounts for approximately 97% of the variation in average task scores. According to the model output, relative noun frequency is most predictive of learner performance with grammatical gender, followed by token grammaticality and noun gender class, whereas noun morphology does not appear to be significantly predictive of performance when all linguistic factors are considered together. A significant multiple linear regression model was also found for the native Spanish speakers that accounts for approximately 82% of the variation in average task scores. However, according to the model output, none of the individual model coefficients reached significance. Therefore, although when considered all together, the linguistic factors analyzed appear to be largely predictive of native speaker performance, none of these factors are *individually* predictive of performance with grammatical gender in Spanish as a native language.

6.2.1 Linguistic factors: possible explanations

We will now consider possible explanations for the finding that learners demonstrate significantly more accurate performance with frequent, grammatical, and masculine noun tokens, a tendency modulated by increasing proficiency, as well as the finding that variation fluctuates with proficiency and per task type. Recall that conscious or explicit experiences with language do not revolve around counting, and therefore, to the extent that language processing is based on frequency and probabilistic knowledge, language learning is largely implicit in nature and language acquisition, representation, and processing are all tuned to varying degrees by frequencies in the input (N.C. Ellis, 2002). The frequency effects uncovered in the present findings corroborate previous research that has demonstrated significant positive correlations between accuracy and frequency (e.g., N.C. Ellis & Schmidt, 1997; Brown, 1973; Larsen-

Freeman, 1978; Hatch & Wagner-Gough, 1976). There is some evidence in previous research for a connection between the regularity of the target item and the relative effect that frequency has on performance such that when a structure is less systematic (i.e., contradicts the canonical pattern), frequency effects become even more pronounced (N.C. Ellis & Schmidt, 1997). However, with the current analysis, we can only draw conclusions regarding the positive effect of transparent noun morphology on accuracy detected in the descriptive results and initial *t*-testing (although not found to be a significant predictor of performance in regression modeling) and the significant effect detected of relative noun frequency in general. Therefore, further re-analysis of the present data would be needed to uncover any frequency-by-regularity interactions. In addition, target language proficiency was found to interact with the relative effect of noun frequency as beginner and intermediate learners were more affected by frequency than their advanced learner and native speaker counterparts, which corroborates previous findings (e.g., N.C. Ellis & Schmidt, 1997; Gass & Mackey, 2002) that have detected a diminishing effect size of frequency with increasing language proficiency. Gass and Mackey (2002) take the interpretation of frequency effects a step further relating the effect of frequency to maturational constraints; they argue that it is possible that the age effects observed in much empirical research actually relate to sensitivity to frequency of input (in implicit learning) and that this may also help to explain differences between native and late nonnative language acquisition. Nonetheless, our findings may contradict this conclusion as both native speaker and late learner performance was found to be significantly predicted by the relative frequency of the target nouns.

With regards to the apparently systematic variation uncovered in both native speakers and adult learners, recall that intra-learner variation is key to understanding the dynamic system of language development (Larsen-Freeman, 2012). Furthermore, there is growing evidence of considerable individual differences in linguistic competence not only in adult learners but also in adult native speakers (Dąbrowska, 2012; Farmer et al., 2012; Hulstijn, 2015). Such native speaker differences are attributable in part to differences in language-related experience, such as level of education and degree and depth of print exposure (i.e., literacy), and also partly attributable to learner-internal factors such as statistical learning abilities, intelligence quotient (IQ), and metalinguistic abilities (Andringa & Dąbrowska, 2019). Our findings corroborate this idea of variation in linguistic competence and metalinguistic abilities in both native and nonnative language users as we uncovered considerable variation at both the inter- and intra-speaker/learner levels and, moreover, we demonstrated how metalinguistic awareness of the target structure is a factor that influences both native speaker and nonnative learner performance. Moreover, metalinguistic awareness appeared to be most influential on performance accuracy when participants had more time to self-pace their performance and thereby tap into their explicit or metalinguistic knowledge, and this was the case for *both* late learners and adult native speakers alike.

Furthermore, recall that errors with grammatical gender can be directly linked to the linguistic variables involved in grammatical gender agreement and assignment, including gender class (masculine/feminine), domain of agreement (article-noun/noun-adjective), and morphology (overt/non-overt) (see: Foote, 2015; Gamboa, 2012: Alarcón, 2011; Montrul et al., 2008). When considering the influence of these linguistic features inherent to grammatical gender, our findings align with previous research showing that native speakers, and not just late learners, also exhibit certain preferences or biases for the masculine gender class, treated as 'default' when assigning gender to determiners of unknown nouns, nonce words, and loanwords (e.g., Eddington, 2002; Eddington & Hualde, 2008; Harris, 1991). Although the present study does not intentionally examine performance with unknown nouns nor nonce or loanwords, (although beginner learners in the study were surely not familiar with all of the target nouns), we add to this body of work with the present findings by instantiating not only a bias for the masculine noun class but also a bias for overtly marked and high-frequency nouns. These biases exist to a much smaller extent in native speakers but are nonetheless present, particularly at the descriptive level of analysis. Therefore, we reaffirm our initial assertion that the use of morphological cues and masculine as default are not phenomena unique to late learners but rather appear to constitute general linguistic processing strategies used by *all* language users.

Our findings on how the linguistic features of grammatical gender modulate late learner and native speaker performance partially corroborate previous findings. For example, some studies (Foote, 2015; Montrul et al., 2008) have found that late learners use noun morphology as a cue when processing and learning grammatical gender since learners are more accurate with nouns that have overt morphology for gender, that is, nouns exhibiting canonical morphology. Our initial descriptive findings confirm that overt noun morphology is associated with relatively higher accuracy scores, although the effect of noun morphology was non-significant in our linear regression modeling to predict performance with the present sample. Therefore, the finding that morphology does play a role in the descriptive findings but ultimately is not a significant predictor of performance may provide some support for the conclusions of Alarcón (2009) who reported that late learners do not appear to process grammatical gender morphologically according to reaction times on a sentence completion task. With regards to the effect of noun gender class, our results corroborate those of Gamboa (2012) and Montrul et al. (2008) who both found that late learners tend to overgeneralize the masculine gender class and erroneously apply it to feminine inanimate nouns. However, our results run contrary to the assertions of López Prego (2015) who found that late learners of Spanish were actually more accurate with feminine nouns. The present study adds to this body of research on how learner performance is affected by the linguistic features of grammatical gender by providing new data on token grammaticality and frequency effects; our findings provide evidence that learners exhibit significantly enhanced accuracy on grammatical and high-frequency tokens, which has not been considered in previous research on the late acquisition of grammatical gender in Spanish. Perhaps the mixed findings uncovered in previous studies on the effect of noun morphology and gender class could be somewhat explained by the grammaticality of the target tokens and by the relative frequency of the nouns included in the target stimuli.

Returning to the concepts of Complexity and Dynamic Systems Theory and different knowledge types and long-term memory systems in language acquisition, recall that treating variation as a dependent response variable allows us to draw some principled conclusions regarding the knowledge source and processing type evident in native speaker and learner performance since explicit knowledge is posited to be much more variable than implicit knowledge (R. Ellis, 2005). Furthermore, native speakers can exhibit sensitivity attuned to distributional patterns of language use (Beatty-Martínez & Dussias, 2019), and, in fact, our linear regression modeling demonstrated that relative noun frequency is an independently predictive factor for both native speaker and late learner performance. Variability has also been observed in native speaker processing of grammatical gender and differential behavior of masculine and feminine gender has been observed in that the masculine is commonly treated as a default form in both native and nonnative performance (e.g., Beatty-Martínez & Dussias, 2019). Our results further confirm that variability is certainly present in both speaker groups, albeit to a larger extent in late learners, and largely runs inverse to developing proficiency level such that as proficiency increases, variation in performance decreases at both the group (standard deviation) and individual (variance) levels. These findings support the assertions of how variation can be interpreted from a Complexity and Dynamic Systems Theory approach in second language acquisition research; as the developing mental grammar system becomes consolidated (i.e., increased proficiency), it approaches an "attractor state" that is more systematic and therefore subject to much less variation (Larsen-Freeman, 2012). Furthermore, our findings demonstrate how variation is conditioned in part by the particular task demands and this is the case across all speaker groups; native speakers and learners of all proficiency levels demonstrate relatively more variation in their performance on self-paced tasks than on speeded tasks. Evidently, tasks which allow for sufficient time to tap into explicit and metalinguistic knowledge stored in declarative memory result in much more variation in performance precisely because explicit knowledge is much less systematic and therefore more variable.

In sum, the finding that learners demonstrate significantly higher accuracy with frequent, grammatical, and masculine noun tokens, modulated by increasing proficiency, as well as the finding that variation fluctuates with proficiency and per task type, can be explained in part by Complexity and Dynamic Systems Theory (Larsen-Freeman, 2012) as well as by morphological cues (Foote, 2015; Montrul et al., 2008) and default processing strategies (Eddington & Hualde, 2008; Harris, 1991).

6.3 Task effects (RQ2)

The second research question examined how the varied task demands (time constraint, stimuli presentation modality) impact both accuracy and variation in performance in native speakers and adult learners. Task effects were expected for both types of task characteristics. With regards to the time constraint task factor, lower accuracy scores and higher variation were expected in adult learners on the speeded tasks (Tasks 2-4). With regards to stimuli presentation modality, lower accuracy and higher variation were expected for the auditory modality as compared to the written modality. Native speakers were also expected to show some task effects in their performance, but to a far lesser extent than adult learners. These predictions were largely corroborated by the results.

Beginner and intermediate learners were clearly affected by task type and performed better on the self-paced tasks and on the written tasks. In contrast, both the native speakers and the advanced proficiency learners actually performed better on the speeded tasks and showed virtually no difference in their accuracy scores between written and auditory tasks. Furthermore, task type appeared to have an impact on variation in performance such that both native speakers and learners were more variable in their responses on self-paced tasks as well as with auditory stimuli tasks, and beginner learner variation was much more affected overall by task type than the other native speaker and learner groups.

To determine to what extent the descriptive findings on task effects could be corroborated by inferential analysis, multiple linear regression models with a confidence interval of 95% were also constructed. For adult Spanish learners, a significant multiple linear regression model was found that accounts for approximately 96% of the variation in average task scores. According to the model output, the time constraint factor was the most powerful predictor of performance in adult learners, followed by the stimuli modality task factor, and when both task factors were considered together, they were highly predictive of performance with grammatical gender in Spanish as a nonnative language. A significant multiple linear regression model was also found for the native Spanish speakers that accounts for approximately 91% of the variation in average task scores. According to the model output, the time constraint task factor was most predictive of performance; however, contrary to adult learners, native speakers performed significantly *better* under a time constraint, and furthermore, task stimuli modality was *not* independently predictive of performance with grammatical gender in Spanish as a native language, contrary to the trend observed in adult learners.

6.3.1 Task factors: possible explanations

We will now consider possible explanations for the findings that: 1. intermediate and beginner adult learners performed better on self-paced and written tasks; 2. that these task effects interacted with proficiency such that advanced learners and native speakers demonstrated enhanced performance on speeded tasks and no significant differences in accuracy according to stimuli modality; and 3. that the time constraint factor was most predictive of performance in all learners and native speakers alike. Recall that there are two main modes of language learning that both occur to some degree in all learners and whose relative proportions in any given learner depend principally on age and context: *implicit* language learning and *explicit* language learning (N.C. Ellis, 2015). Therefore, acquiring language competence takes place through exposure to comprehensible input in two main modes: *implicitly* via the nonconscious and automatic abstraction of structural patterns and *explicitly* through selective learning in which the learner seeks out patterns and tests hypotheses. Differences in performance conditioned by varying task demands detected in the present sample could be largely explained by the use of different language processing modalities such that explicit knowledge about language is conditioned on the self-paced tasks and on the metalinguistic awareness task, whereas implicit linguistic competence is conditioned on the speeded tasks, encouraging participants to respond to task prompts according to intuition under a time constraint. In other words, we can relate the observed task effects to varying degrees of implicit competence and explicit knowledge about language that are posited to be represented to varying proportions in native speakers and adult learners of varying proficiency levels.

Although implicit learning is the norm in native acquisition, the overwhelming consensus in late acquisition research is that nonnative adult acquisition by implicit means alone is limited in its success (e.g., Schmidt, 1990; Lightbown et al., 1993). Perdue (1993) and Klein (1998), for example, show how years of language input exposure can fail to become target-like linguistic competence. Schmidt (1990, 2001) explains that implicit learning mechanisms such as statistical tallying do not take place in cases where linguistic form lacks perceptual salience. Therefore, in such cases, additional attention is necessary for the relevant associations to be learned and this can be achieved through form-focused instruction that recruits learners' explicit conscious processing (N.C. Ellis, 2005). Why are adult learners less able to use implicit learning mechanisms? N.C. Ellis (2005) explains that years of learned attention during native language and general cognitive development limit the potential for implicit learning in subsequent language acquisition. Therefore, explicit learning, and at least some explicit instruction, seems to be necessary to reach target-like norms in late nonnative language development. Nonetheless, language development—both native and nonnative—is a complex, dynamic, and adaptive system in which interactions between the implicit and the explicit constitute constant currents of change and development (Beckner et al., 2009; N.C. Ellis, 2007; N.C. Ellis & Larsen-Freeman, 2006, 2009). As such, there has been persistent debate on how implicit and explicit knowledge systems may be related (see: Bialystok, 1994; R. Ellis, 1993, 2005; Krashen, 1981; Paradis, 1994; Hulstijn, 2002) and three different theoretical perspectives have emerged, including the *non-interface position*, the *strong interface position*, and the *weak interface position*.

Although the non-interface position posits that implicit and explicit language knowledge involve different acquisitional mechanisms (Hulstijn, 2002; Krashen, 1981), are stored in different parts of the brain (Paradis, 1994), and are accessed for performance by different processes, either automatic or controlled (R. Ellis, 1993), the other two positions recognize an important connection between the two knowledge systems. Explicit knowledge can contribute indirectly to the acquisition of implicit linguistic knowledge by promoting some of the acquisition processes believed to be responsible. For example, N.C. Ellis (1994) suggests that declarative 'rules' can have an influence on perception by making relevant features salient and thus enabling learners to notice them and to notice the gap between the input and their existing linguistic competence. Furthermore, Schmidt and Frota (1986) have argued that learners can use their explicit language knowledge to produce output that then serves as 'auto-input' to their implicit learning mechanisms. Different iterations of the implicit-explicit interface debate recognize that all language users possess both implicit and explicit linguistic knowledge, albeit in different proportions, and that both knowledge types have the potential to develop in tandem and simultaneously—whether one is transformed into the other or not. Independently of which interface position you adopt, there is wide acceptance that explicit knowledge can contribute to performance such that different performance tasks are likely to induce learners to draw differentially on their implicit and explicit knowledge. The task effects detected in the present study corroborate the findings of Bialystok (1982), for example, in which a writing task induced learners to draw more extensively on their explicit/declarative knowledge than a task that solicited unplanned oral communication, in line with the speeded vs. self-paced and auditory vs. written dichotomy of task effects observed in the present findings. Furthermore, the finding that both native speakers and nonnative learners alike were significantly affected by task effects—in particular the time constraint component (although the direction of the effect varied)—illustrates that both implicit and explicit knowledge and memory systems are at play to varying degrees in different groups of language users. Moreover, the present findings provide evidence of how developing proficiency interacts with this effect such that implicit language processing is favored at advanced and native levels of proficiency—contributing to a larger extent to average accuracy scores—while explicit language processing is favored at beginner levels of proficiency leading to diminished accuracy on tasks that require more extensive use of implicit linguistic competence (e.g., speeded tasks).

Recall that Ullman's (2001, 2004) Declarative-Procedural (DP) model formalizes the distinction between the declarative and procedural long-term memory systems for language development. In contrast to L1 grammar, both lexical development as well as early L2 grammar development are argued to rely more on declarative memory in which only at higher proficiency levels can grammar learning take place in the procedural system (Hamrick, 2015); Morgan-Short et al., 2014). Therefore, individual differences in declarative and procedural memory abilities have been shown to correlate with L2 learning (Carpenter, 2008; Morgan-Short et al., 2015). This distinction between declarative and procedural long-term memory systems is operationalized via online vs. offline language processing; recall that *online* language processing requires the use of *implicit* knowledge stored in the *procedural* memory system, while *offline* processing allows the language user time to tap into their *explicit* language knowledge stored in the *declarative* memory system. The task effects finding uncovered in the present study corroborates this distinction per proficiency level in which the tasks designed to solicit online language processing accessing procedural long-term memory under a time constraint (i.e., the speeded tasks) led to lower accuracy scores in beginner and intermediate learners as compared to offline tasks in which participants had sufficient time to access their explicit language knowledge stored in declarative memory (i.e., self-paced tasks). In contrast, both native speakers and advanced learners actually performed better on online tasks in which they were conditioned to tap into implicit linguistic competence stored in procedural memory. In sum, the present findings align with a declarative-procedural memory distinction via accuracy differences during offline as compared to online language performance tasks, and furthermore, the interaction detected with proficiency level indicates that the relative balance of access to declarative and procedural memory systems changes with proficiency such that native speakers and high-proficiency nonnative learners access their procedural memory to a larger degree and with greater success than their lower-proficiency peers. However, we must be cautious to not overinterpret the task effects findings uncovered in the present study. Principally, there can be no guarantee that 'task-as-workplan' will directly and consistently correspond to the 'task-as-process', as argued by some researchers (e.g., Breen, 1989; Coughlan & Duff, 1994). Therefore, tests designed to operationalize this distinction are expected to *predispose* learners to access one or the other knowledge type, but only in a *probabilistic* manner.

Regarding the native vs. nonnative dichotomy and the potential for incomplete acquisition in late learners, we have established that different performance tasks are believed to induce late learners to draw differentially on their implicit and explicit knowledge (Bialystok, 1982), yet this is not, however, what is expected in native speakers. In order to separate the effects of task-related factors from effects that are due to incomplete acquisition of the target language, it is essential to examine how task manipulations influence the linguistic performance of both native and nonnative speakers—something that few studies have endeavored to do (e.g., Michel et al., 2019; Foster & Tavakoli, 2009). One of the few studies that compares native speakers to nonnative learners with regards to task effects is Michel et al. (2019) who found that the effect of cognitive task demands on syntactic complexity and accuracy varied according to task type and speaker status such that L2 users produced more subordinate clauses on the more cognitively demanding decision-making task yet for the L1 group, the particular demands of the task had no effect on complexity. As previously discussed, the findings of the present study contradict this as both native speakers and late learners were significantly affected by task effects, in particular the time constraint component, although the direction of this effect was varied such that native speaker and advanced learner performance was enhanced by a time constraint and intermediate and beginner learner performance deteriorated under a time constraint. Nonetheless, the stimuli modality factor was a significant predictor of performance in late learners only. Why might less proficient late learners benefit from the ability to self-pace their performance while advanced proficiency learners and native speakers do not? A possible explanation lies in how information is processed. Information Processing Theory claims that humans possess a limited processing capacity and, as a result, are not able to attend fully to all aspects of a task simultaneously

(Anderson, 1995; Newell & Simon, 1972). However, when language learners have the opportunity to plan linguistic and propositional content before and/or during language performance, they can compensate for these processing limitations and, as a result, the quality of their linguistic output in enhanced (Skehan, 1996). In this sense, planning helps learners to access linguistic material from memory more easily, particularly items stored in declarative memory that by nature require greater working memory to retrieve. Therefore, if an adult learner is at a lower level of proficiency in the language, they likely depend to a larger extent on declarative memory stores to compensate for a (temporary) deficit in implicit linguistic competence and, therefore, benefit from the opportunity to tap into this declarative knowledge source whereas automatic and fluent language processing is less effortful in the absence of such a deficit in implicit linguistic competence, as is the case in advanced learners and native speakers.

The stimuli modality effects uncovered in the late learner group can be partially explained in psychological terms. Psychological research on modality differences suggests that verbal material presented aurally and visually is processed in different parts of the memory system and by different mechanisms (Penny, 1989). Furthermore, language studies suggest that the written modality enables more elaborate (e.g., Vasylets et al., 2017) and more accurate (e.g., Kormos, 2014) language production than the aural modality and these claims have been corroborated by empirical language acquisition research that has shown higher accuracy rates (Bialystok, 1979, 1982; Johnson, 1992) for written grammaticality judgement tasks (GJT) when compared to results obtained from GJTs presented in the aural modality (e.g., Haig, 1991; Johnson, 1992). Murphy (1997) explains this dichotomy in terms of the perceived “burdens of auditory processing” (Murphy, 1997, p. 55) and the ease of visual processing due to heightened cognitive demands imposed on language users when processing auditory stimuli. The findings of the present study corroborate previous findings from SLA literature that suggest that the untimed nature of written input allows for better information uptake than auditory input does (Bialystok, 1997, 1982; Johnson, 1992; Murphy, 1997) as intermediate and beginner late learners demonstrated higher accuracy rates with the written modality. Furthermore, speaker status as native or nonnative also variably interacts with the effect of stimulus modality (e.g., Murphy, 1997). Our findings partially corroborate Murphy (1997) who found that although differences per modality were observed for both native and nonnative speaker groups, auditory processing seemed to produce a greater obstacle for L2 learners to overcome than for native speakers. In the present findings, the late learners were more accurate with written as opposed to auditory stimuli tasks; however, this dichotomy in performance per modality type was not found to be the case in the native speaker group nor in the advanced proficiency adult learners. Therefore, stimuli modality effects differentiate not just early and late language acquisition, but more precisely align with developing proficiency: as proficiency increases, the modality of the stimuli becomes less important as late proficient learners no longer rely on the scaffolding effect of written input.

In sum, the finding that intermediate and beginner adult learners performed better on self-paced and written tasks and that these task effects interacted with proficiency such that advanced learners and native speakers demonstrated enhanced performance on speeded tasks and no differences in accuracy according to stimuli modality, combined with the finding that the time constraint factor was most predictive of performance in all learners and native speakers alike, can be explained in part by the theoretical distinction between implicit and explicit learning mechanisms and knowledge types (N.C. Ellis, 2005), the Declarative Procedural Model (Ullman, 2001, 2004) for long-term memory systems, and in part by Information Processing Theory (Anderson, 1995).

6.4 Individual factors (RQ3)

The third research question examined to what extent individual factors, including self-reported and tested Spanish proficiency, metacognitive and metalinguistic awareness, motivation, attitudes about the target language and community, Spanish language use, and linguistic repertoire may predict or account for accuracy and variation in learner performance. Furthermore, the interaction between learner factors and task type was explored to determine to what extent individual factors account for performance on varied task demands. It was predicted that all individual variables analyzed would have some effect on learner performance, albeit with varying effect sizes. It was also predicted that although native speakers would show some degree of variation across the individual variables, these individual factors were not expected to have a significant impact on native speaker performance. Furthermore, recall from the Methodology section, that data was not collected on native speaker motivation nor metacognitive awareness. For late learners, proficiency was predicted to have the largest effect on performance, and some significant interactions between variables were also expected, including between metacognitive awareness and metalinguistic awareness, between metalinguistic awareness and linguistic repertoire, and between motivation and attitudes. With regards to the interplay between the individual variables and the task factors, it was predicted that as proficiency increases, the effect of the time constraint task factor would decrease, meaning that advanced learners would demonstrate less of a difference in accuracy scores between task types than intermediate and beginner learners. Metacognitive awareness was also predicted to be more associated with self-paced tasks such that learners who scored higher on the metacognitive awareness inventory would also show significantly higher accuracy scores on the self-paced tasks than on the speeded tasks, and that this effect would be more pronounced than for less metacognitively aware learners. These predictions were somewhat corroborated by the results.

6.4.1 Individual factors: Spanish proficiency

Tested Spanish proficiency treated as a categorical independent variable (i.e., beginner, intermediate, advanced) appeared to have an effect on both accuracy and variation across learner proficiency groups in which high proficiency learners patterned very closely to the native Spanish speakers. Furthermore, higher proficiency level was associated with higher accuracy, particularly with ungrammatical feminine noun tokens and on auditory stimuli tasks, and group variation (SD) in performance also decreased with increasing proficiency level. It is important to note that scores on ungrammatical feminine nouns are likely the most indicative of acquisition of the grammatical gender target structure since we know from previous empirical findings that the masculine gender class tends to be treated as the 'default', both in native and nonnative processing of Spanish gender (e.g., Eddington, 2002; Eddington & Hualde, 2008; Harris, 1991), and our present data seem to suggest that 'grammatical' is also treated as a default response by late learners. Therefore, to the extent that late learners are accurate with ungrammatical feminine nouns, they are more likely drawing upon their acquired linguistic knowledge (either implicit or explicit or both) of grammatical gender as they are not able to rely on the prototypical default response with these tokens. Likewise, to the extent that late learners make errors with ungrammatical feminine nouns, they likely exhibit a deficit in their competence with grammatical gender.

The different dimensions of Spanish proficiency, including self-reported, vocabulary scores, grammar scores, and total tested Spanish proficiency, were also treated as a continuous independent variable for analysis and the output of a Pearson's *r* correlation matrix indicated that Spanish proficiency was significantly and positively correlated with performance in adult learners. This association was most pronounced on speeded and auditory tasks and seemed to be more associated with tested vocabulary scores than tested grammar scores. Furthermore, intra-learner variance was found to be highly and negatively correlated with all measures of Spanish proficiency, but again showed the strongest negative correlation with tested vocabulary scores in adult learners. Native Spanish speakers also demonstrated some moderate to strong correlations between their Spanish proficiency scores and scores on the different task types; however, correlations in native speakers were larger and more significant with self-reported Spanish proficiency than with tested Spanish proficiency and notably no significant correlations were detected between tested grammar scores and task scores in native speakers. Therefore, it seems that native speakers are more accurate at self-reporting their fluency in Spanish as their self-report appears to be more closely correlated with their actual performance with a particular grammar structure (i.e., grammatical gender) than adult learners, who show the strongest correlation between their tested vocabulary scores and their performance.

6.4.2 Individual factors: Metacognitive awareness

Metacognitive awareness treated as a categorical independent variable (i.e., more aware, less aware), appeared to have an effect such that more metacognitively aware learners demonstrated higher average accuracy scores, and this advantage was particularly evident during self-paced tasks and on the more difficult ungrammatical feminine noun token type. However, being more metacognitively aware was *not* associated with lower group variation (SD) in average accuracy scores.

Metacognitive awareness level and type were also examined as continuous independent variables in adult learners of Spanish. The output of a Pearson's *r* correlation matrix between the different components of metacognitive awareness, including *knowledge about cognition* and *regulation of cognition*, and average scores on different task types, indicated a moderate significant correlation between metacognitive awareness and average speeded task scores as well as average auditory task scores. Nonetheless, the effect of metacognitive awareness appeared to be limited to *knowledge about cognition* and not the *regulation of cognition* and was only detectable with certain task types (speeded and auditory) but was notably *not* significantly correlated with task scores overall.

6.4.3 Individual factors: motivational orientation

Motivation was first treated as a categorical independent variable (i.e., positive, negative, neutral). Results indicated that learners with a more positive motivational orientation consistently demonstrated higher average accuracy rates than their peers with a more negative motivational orientation, and this advantage was particularly evident on speeded and auditory tasks. Nonetheless, the observed difference between motivational orientation groups was binary in nature in which clear distinctions were only observable between positive and negative motivational orientation groups, whereas the effect of a more neutral motivational orientation was less evident. Furthermore, motivational orientation does not appear to have an effect on group variation (SD) in learner performance.

Motivational orientation was also examined as a continuous independent variable in adult learners of Spanish. The output of a Pearson's *r* correlation matrix between the different components of motivational orientation and average accuracy scores on different task types uncovered moderate to strong positive correlations, especially with speeded and auditory tasks, in addition to significant negative correlations with intra-learner variance. With regards to the specific components of motivational orientation, the *ideal L2 self*, the *ought-to L2 self*, the *experiential* component of motivation, as well as *overall motivation* scores were all significantly and positively correlated with average accuracy scores, while the *behavioral* component of motivation was notably *not* correlated with learner performance on any task.

6.4.4 Individual factors: Spanish use

The effect of Spanish language use was also examined in both native Spanish speakers and Spanish learner groups. Treated as a categorical independent variable (i.e., low, moderate, high use), results indicated that learners who reported moderate use of Spanish in a typical week clearly outperformed their peers with low reported use of Spanish, particularly with the more difficult ungrammatical feminine noun token type and on the auditory task type. In contrast, native speaker performance appeared to be largely unaffected by average weekly use of Spanish; however, some advantage was observed in performance with ungrammatical feminine noun tokens and also on self-paced tasks for those native speakers who reported high weekly use of Spanish.

Average weekly Spanish language use was also examined as a continuous independent variable in both native Spanish speakers and adult Spanish learners. The output of a Pearson's *r* correlation matrix indicated that average weekly use of Spanish is moderately and significantly correlated with accuracy on auditory, written, and speeded tasks and significantly negatively correlated with intra-learner variance. Therefore, the benefit of increased target language use was *not* found to be associated with tasks in which learners can take their time (i.e., self-paced) nor was it linked to their performance with the more difficult token types (i.e., ungrammatical feminine nouns) believed to be more indicative of acquisition of grammatical gender, independent from the prototypical default response. Native speakers, in contrast, demonstrated a significant moderate positive correlation between average reported weekly use of Spanish and average accuracy scores on ungrammatical feminine nouns *only*; no other significant correlations were detected with average Spanish use in native speakers.

6.4.5 Individual factors: Metalinguistic awareness

The effect of metalinguistic awareness was also examined in both native Spanish speakers and adult Spanish learner groups. Treated as a categorical variable (i.e., more aware, less aware), it was found that being more explicitly aware of the grammatical structure being tested (i.e., metalinguistic awareness) affords certain advantage in adult learners and this advantage was found to be most pronounced on the more difficult token type of ungrammatical feminine target nouns, posited to be most indicative of acquisition since this token type has been shown to not be as susceptible to a default response. Moreover, *more aware* learners as a group demonstrated less variation (SD) in their accuracy scores than their *less aware* learner peer group. Surprisingly, native Spanish speakers' performance was also found to be affected by their metalinguistic awareness of grammatical gender, yet to a lesser extent than their adult learner counterparts, and this metalinguistic awareness effect detected in native speakers was most pronounced on self-paced tasks and on the more difficult ungrammatical feminine noun token type. Similarly to the adult learner group, *more aware* native speakers also demonstrated relatively less variation (SD) as a group than their *less aware* native speaker peers.

Pearson's *r* correlation matrices between the different forms of metalinguistic awareness (overt, conditioned, total) and average accuracy scores uncovered moderate to strong significant positive correlations between accuracy scores on all task types and all forms of metalinguistic awareness scores in adult learners, yet this association was found to be most pronounced with self-paced and written tasks. Furthermore, a significant and moderate negative correlation was found between average intra-learner variance in performance and both overt and total metalinguistic awareness. Native speakers' accuracy scores were found to be correlated with their level of metalinguistic awareness about grammatical gender as the target structure, yet to a lesser extent than their adult learner peers. Furthermore, this correlation was found to be strongest with *conditioned* awareness rather than overt or total metalinguistic awareness, and average intra-speaker variance was *not* significantly correlated with metalinguistic awareness in native Spanish speakers, contrary to their adult learner peers.

6.4.6 Individual factors: attitudes

Average accuracy scores in adult learners were compared according to self-reported attitudes towards the Spanish language and the target language community. Recall that all native Spanish speakers responded positively to the Spanish language attitudes portion of the Language Learner Profile Questionnaire, and therefore the effect of their reported attitudes on their performance was not considered further as an independent explanatory variable to predict performance with grammatical gender. When the attitudes individual factor was treated as a categorical variable (i.e., positive, negative, neutral attitude), it was found that adult learners with more *positive* attitudes toward the target language and target language community demonstrated consistently higher accuracy rates than their *negative* attitude learner peers, and this effect was found to be particularly pronounced during the speeded and auditory task types as well as with the more difficult ungrammatical feminine noun tokens, more indicative of acquisition as this token type is not associated with a default response. Nonetheless, similarly to the effect of motivation, the effect of attitudes was found to be binary in nature in which the clearest distinctions were observable between positive and negative attitude groups, yet the effect of a more neutral attitude was less clear.

The output of a Pearson's *r* correlation matrix between average Spanish attitude scores and average accuracy scores on each task type demonstrated that in adult learners more positive attitudes about the Spanish language and target language community were significantly and positively correlated with accuracy scores and negatively correlated with intra-learner variance, and this was an association found to be most pronounced with auditory and speeded task types, those that are posited to tap more into implicit linguistic competence in procedural memory.

6.4.7 Individual factors: prior linguistic repertoire

The effect of prior linguistic repertoire in terms of the presence of grammatical gender was also investigated in late/adult learners of Spanish. Recall that native Spanish speakers were excluded from the analysis of prior linguistic repertoire since their native language (Spanish) features a grammatical gender system, and there is no principled way to tease apart the effect of a gendered L1 from a gendered L*n* in one's linguistic repertoire, despite the fact that half of the native Spanish speaker participants self-reported as 'multilingual'. When prior linguistic repertoire was treated as a categorical variable (i.e., grammatical gender 'present' or 'absent'), the group of multilingual learners with grammatical gender already 'present' outperformed their multilingual learner peers with grammatical gender 'absent' in their prior linguistic repertoire. The positive effect of prior linguistic repertoire was found to be most pronounced on speeded and auditory task types—posited to tap more into implicit linguistic competence—and on the more difficult ungrammatical feminine noun token type that is likely more indicative of real acquisition of the target structure since it is not associated, on average, with a default response. Nonetheless, in terms of group variation (SD), prior linguistic repertoire did not appear to have a differentiating effect between multilingual learners with and without grammatical gender already present.

6.4.8 Individual factors: combined effects

To investigate how individual differences together impact performance with grammatical gender in both native Spanish speakers and adult Spanish learners, multiple linear regression models were constructed at a confidence interval of 95%. A significant model was found that accounts for approximately 67% of the variation in accuracy scores in adult Spanish learners. The respective model coefficients indicated that total Spanish proficiency is most predictive of learner performance with grammatical gender, followed by metalinguistic awareness of the target structure and motivation to learn the language, whereas all other individual factors were *not* found to be *independently* predictive of adult learner performance. Contrary to adult learners, a significant model was *not* found for native Spanish speakers, indicating that the individual factors of proficiency, metalinguistic awareness, attitudes, and language use are not significantly predictive of performance in native Spanish speakers.

6.4.9 Individual factors: inter-factor correlations

Finally, recall that in order to examine more closely the nature of individual learner differences and how different predictors of learner performance may be intertwined, a Pearson's *r* correlation matrix was run between these factors. The strongest positive correlations were observed between Spanish attitudes and motivation, Spanish proficiency and Spanish attitudes, metacognitive awareness and motivation, and between Spanish proficiency and motivation. These inter-factor correlations indicate that more motivated learners tend to have more positive attitudes about the target language community—or learners with more positive attitudes tend to be more motivated to learn the language—and also tend to be more aware of their own language learning process (i.e., metacognitive awareness). Moreover, more proficient learners tend to have more positive attitudes toward the target language community and demonstrate higher levels of motivation to learn the language. Spanish use was also found to be moderately correlated with proficiency, attitudes, and motivation as was metalinguistic awareness with proficiency and motivation thereby indicating that as target language proficiency increases, use of the language, attitudes about the language, and motivation to learn the language also increase, and explicit knowledge about the language (i.e., metalinguistic awareness) is enhanced. Interestingly, regarding learner awareness, metacognitive awareness and metalinguistic awareness were *not* found to be significantly correlated, indicating that being explicitly aware of the target structure and being explicitly aware of the language learning process and associated strategies do not tend to co-occur. Therefore, there does not appear to be a general construct of 'learner awareness' as these two domains are clearly distinguishable from one another. Furthermore, metacognitive awareness was not found to be significantly correlated with Spanish proficiency, perhaps indicating that being explicitly aware of one's individual learning process and strategies is not necessary to develop proficiency in the target language—a finding which may have important implications for how we approach the teaching of learning strategies in language instruction, although this would require further research to determine if perhaps a certain component of metacognitive awareness (i.e., knowledge about or regulation of cognition) may be variably associated with developing proficiency, perhaps at lower proficiency levels. Nonetheless, many of the analyzed individual factors are clearly intertwined and may develop together in unison, affecting learner performance in the target language at any point in time. However, it is important to note that from this data alone we cannot conclude what the direction of influence may be between these individual learner factors. In other words, it is not clear whether having increased proficiency in the language leads to increased motivation, more positive attitudes, and increased use of the language, or whether learners may start off with more positive attitudes and higher motivation levels which drives them to use the language more in their daily lives and thereby increase their overall proficiency level. Further research with a longitudinal approach that follows learners' changing individual characteristics as their proficiency in the target language develops may elucidate these preliminary findings regarding the direction of influence of individual learner factors.

6.4.10 Individual factors: possible explanations

We will now consider possible explanations for the finding that proficiency, metalinguistic awareness, and motivation are all factors that variably modulate late learner performance. We will leave aside the other individual factors which did not produce a significant effect. Recall that patterns seeming to support the existence of a critical or sensitive period for language acquisition have also been explained to some degree by a range of other factors, such as motivations and attitudes (Birdsong, 2005, 2006; Birdsong & Vanhove, 2016; Singleton & Muñoz, 2011). In other words, past the hypothesized critical period for language learning, age is no longer such a deterministic factor and other individual learner variables take on greater importance in predicting performance and ultimate attainment.

First, we will attempt to explain the role of proficiency as an individual factor to predict late learner performance based on how language processing is altered with developing proficiency.

Recall that cognitive and neurological research have uncovered two long-term memory systems. *Declarative memory* supports the learning of general facts and knowledge and consists of explicit knowledge whereas *procedural memory* supports motor and sequential skills and consists of implicit knowledge (Knowlton & Moody, 2008). In terms of language functions, declarative memory is posited to be involved in the acquisition of the mental lexicon whereas procedural memory is believed to be involved in the acquisition of the mental grammar, as posited in Paradis' Neurolinguistic Theory of Bilingualism (2004, 2009)[3], and the declarative-procedural long-term memory distinction is formalized for language learning in Ullman's Declarative-Procedural (DP) Model (2004, 2005). The findings of the present study largely corroborate this distinction as a stronger correlation was detected between tested vocabulary scores and accuracy scores than between tested grammar scores and accuracy scores in adult/late learners, possibly indicating that learner performance is more dependent on and supported by explicit knowledge stored in the mental lexicon of declarative memory. Moreover, adult learners as a group outperformed their native speaker peers on the metalinguistic awareness survey, which is an indicator of one's explicit knowledge about the language structure in question. Therefore, late learner performance is not only affected by developing proficiency in the language but also crucially by the source of linguistic knowledge that is fueling linguistic behavior.

Motivation, in addition to aptitude, has served as the most consistent predictor of L2 learner success, producing correlations with language achievement that range between .20 and .60, with a median value around .40. Aside from age of onset (AO) and aptitude, no other potential predictors of L2 learning success have consistently achieved such levels (Dörnyei & Skehan, 2008). The findings of the present study corroborate previous motivation research with a correlation coefficient detected between motivation and accuracy in performance of .36 to .56. Gardner (1985) argued that an integrative motivation orientation reflects a positive disposition toward the target language community and a desire to interact with and become similar to valued members of that community. Studies have suggested that learners who have a strong desire to integrate into the target language community are both more motivated overall and demonstrate

[3] However, when it comes to grammatical gender, this distinction may be more complex as noun gender assignment is likely encoded as an inherent property along with the lexical entry in the mental lexicon as a sort of "word grammar". Therefore, in the morphosyntactic domain of grammatical gender, mental lexicon and mental grammar likely overlap. See R. Hudson (2007, 2021) for a discussion on word grammar.

higher language attainment than learners who report being more instrumentally oriented, such as for reasons of academic or career advancement (e.g., Gardner & Lambert, 1972). The findings of this study largely corroborate previous motivation research in terms of both achievement and the connection with a desire to integrate in the target language community as a strong significant correlation of .55 was detected between motivational orientation and attitudinal scores, indicating that learners who desire to be perceived as and use the language like a native speaker tend to also be more motivated to learn the language, which intuitively makes sense. However, further research that includes interviews with learners about their specific motivations and attitudes would elucidate the exact nature of this connection. Nonetheless, an integrative motivational orientation may have limited scope and relevance in the instructed learning context outside of immigration and study abroad where such social integration in the target language community is virtually impossible (Au, 1988; Crookes & Schmidt, 1991; Oxford & Shearin, 1994; Leaver, 2003). Moreover, motivation is not, in reality, a static trait but considerably dynamic as it does not remain constant but rather is associated with evolving mental processes, characterized by continuous reappraisal and balancing of the various internal and external influences to which individuals are exposed. Therefore, to further elucidate both the connection uncovered in the present study between a positive motivational orientation and enhanced performance, longitudinal studies are needed to track how adult learners' motivation changes over time with evolving proficiency in the language, and vice-versa.

In order to address the inherent limitations of Gardner's original conceptualization of integrativeness as a motivational desire to integrate into the target language community, Dörnyei (2005, 2009) developed an updated framework, the L2 Motivational Self System (L2MSS), drawing on the psychological theory of "possible selves" (Markus & Nurius, 1986, 1987). Dörnyei and Csizér (2002) reasoned that the process of identification underlying the concept of integrativeness might be better explained as an internal process of identification with a projected future image (i.e., "Ideal L2 Self") within the person's self-concept, rather than identification with an external reference group such as the target language community. The *ideal L2 Self* concerns a desirable self-image of the kind of L2 user one would ideally like to be in the future such that if a discrepancy is observed between the ideal L2 self and one's current state, the individual may be motivated to learn a new language or further develop their proficiency in an existing one. This subcomponent of Dörnyei's L2MSS clearly aligns with and provides a theoretical interpretation of our findings since stronger correlations were detected between the Ideal L2 Self component of motivation and overall accuracy scores, particularly on the speeded and auditory tasks, not only identifying an "ideal L2 self" as the major driving factor of the effect of motivation on performance but also pointing to the nature of the motivational effect as one that potentially impacts implicit linguistic competence more, conditioned during speeded tasks and auditory tasks in which language users must tap into their implicit knowledge in procedural long-term memory. To further explain this connection between motivation and enhanced performance with online language processing tasks, we can refer back to Ehrman's (1996) conceptual framework around the idea of "deep processing", which is an active process of making associations with material that is already familiar and examining interrelationships within the new material. Surface processing, in contrast, occurs when there is no emotional investment, whereas deep processing involves a more profound and integrated interpretation and processing of new material. In this way, processing type is interrelated to some extent with learning motivations and attitudes in the sense that if one feels emotionally engaged with the language and language community, this may fundamentally alter the way the individual processes the target language input thereby leading to deeper, and subsequently more effective, processing. This connection between a motivational orientation grounded in an emotional investment in the target language community and deep processing of language input explains why the effect of a positive motivational orientation seems to be more pronounced with tasks that condition the use of implicit linguistic competence during a time constraint.

Regarding the effect of metalinguistic awareness on language performance, recall that Paradis' Neurolinguistic Theory of Bilingualism (2004, 2009) posits metalinguistic knowledge as one of the four cerebral mechanisms involved in the acquisition and use of language, both native and nonnative, along with implicit linguistic competence, pragmatics, and motivation. L1 children and early L2 learners engage almost solely in incidental acquisition, which arises through procedural memory and leads to implicit competence of linguistic intuition (Paradis, 2004). Adolescents and adult L2 learners, on the other hand, can no longer build procedural representations to the same extent as children and therefore learn the L2 intentionally (to a larger extent) while relying on declarative memory, thereby leading to explicit competence of metalinguistic knowledge and awareness (Paradis, 2004, 2009). To the extent that nonnative speakers have gaps in their implicit linguistic competence, they will compensate by relying more extensively on metalinguistic knowledge (Paradis, 2004). Furthermore, there are considerable individual differences in both nonnative and native language users' linguistic competence (Dąbrowska, 2012; Farmer et al., 2012; Hulstijn, 2015) that are partly attributable to differences in learner-internal factors such a metalinguistic ability (Paradis, 2009). For example, Brooks et al. (2017) examined adult learner performance with Russian gender and case-marking patterns and found that even under implicit learning conditions, individual differences stemmed from explicit metalinguistic awareness of the underlying grammar. Our finding that metalinguistic awareness was highly and significantly correlated with both adult learner and native speaker performance with grammatical gender and that this correlation was strongest on tasks that condition explicit linguistic knowledge (i.e., self-paced tasks), combined with the fact that metalinguistic awareness was identified as an individually significant predictor of learner performance in multiple linear regression modeling, corroborates previous research and theory regarding the role of metalinguistic awareness for explicit language knowledge. Furthermore, a moderate and positive correlation between metalinguistic awareness and proficiency was detected in late learners, suggesting that global proficiency is fueled to some extent by explicit awareness of grammar. Nonetheless, from the present analysis, we cannot yet draw conclusions about the extent to which metalinguistic awareness may compensate for deficiencies in implicit linguistic competence as further analysis is needed to compare the effect of metalinguistic awareness on the performance of learners at varying levels of proficiency. Perhaps metalinguistic awareness is more important at beginner and intermediate levels of proficiency to scaffold language processing when there is a representational deficit in the developing mental grammar of the adult learner, but as global proficiency increases, the need to scaffold linguistic performance with explicit and metalinguistic knowledge subsides.

In sum, the finding that Spanish proficiency, metalinguistic awareness, and motivation are significant predictors of late learner performance and that many of these and other individual factors are positively correlated and therefore likely develop in unison can be explained in part by the Neurolinguistic Theory of Bilingualism (Paradis, 2004, 2009), by the concept of "deep processing" and its relationship to learning motivations and attitudes (Ehrman, 1996), the theory of integrative motivation (Gardner, 1985), the construct of the L2 Motivational Self System (Dörnyei, 2005, 2009), and in part by the psychological theory of "possible selves" (Markus & Nurius, 1986, 1987).

6.5 Multilingual effect (RQ4)

The final research question examined whether or not participants who reported proficiency in additional languages (i.e., "multilingual") would also demonstrate higher accuracy scores than those participants who reported exclusive knowledge of English and Spanish. Furthermore, the possible nature of a multilingual effect was considered by comparing multilingual learners with and without grammatical gender and binary grammatical gender already present in their multilingual linguistic repertoire to determine to what extent the other language typology may contribute to or explain any observed multilingual effect. It was predicted that learners who report knowledge of at least one additional language would demonstrate higher accuracy rates than their exclusively bilingual (English-Spanish) peers and that this effect would be even more pronounced for learners who know more than three languages (i.e., more than one additional reported language). Furthermore, it was predicted that having competence in an additional gendered language would *not* have a significant effect that is independent of the general effect of multilingualism, meaning that multilingual learners with grammatical gender already present in their linguistic repertoire would *not* demonstrate significantly higher accuracy scores than their multilingual peers who do not report knowledge of an additional gendered language. Finally, multilingualism was predicted to have no effect on native speaker performance. These predictions were somewhat corroborated by the results.

6.5.1 Multilingual effect: general advantage

The effect of being multilingual when learning Spanish as an additional language was examined in adult learners and was also considered in native Spanish speakers as a point of comparison. Results indicated that for an adult learner of Spanish, being multilingual affords certain advantage as it is associated with higher accuracy rates across all measures, particularly with the more difficult ungrammatical feminine noun token type (i.e., more indicative of acquisition as this token type is not as susceptible to default processing) and is also associated with higher scores on the more cognitively demanding auditory and speeded task types—posited to exploit to a larger extent implicit linguistic competence. In contrast, for native Spanish speakers, the effect of being multilingual was rather minimal and only afforded a slight advantage on self-paced task scores—the task type with which native speakers on average struggled the most—yet was associated with an even slighter disadvantage on other measures of accuracy and was further associated with greater group variation (SD) in performance. Therefore, it appears that being multilingual may *disrupt* native speaker performance, particularly with the more difficult ungrammatical feminine noun token type, as there may be confusion generated from knowing the gender assignment of nouns in other gendered languages in the multilingual native speaker's linguistic repertoire. Nonetheless, when given more time to contemplate a response, such as on the self-paced task type, multilingual native Spanish speakers might be more able to tap into their declarative linguistic knowledge store that has been deepened through their multilingual language experience after having fully acquired their native language (Spanish).

Further analysis and more data is needed from multilingual native speakers in order to explore *why* being multilingual as a native speaker affords an advantage on self-paced tasks but leads to a relative disadvantage for accuracy on speeded tasks and on more difficult token types. Perhaps there is interference occurring between the gender assignment of feminine nouns in Spanish and the equivalent noun in another gendered language in their linguistic repertoire and perhaps this potential interference generated by cross-language noun gender assignment asymmetry is mitigated to a certain degree when multilingual native Spanish speakers are provided with sufficient time to consciously analyze. The fact that much more group variation (SD) was observed with multilingual native speakers' accuracy scores on ungrammatical feminine noun tokens than on other task measures corroborates this idea of a system in flux, subject to potential interference from competing gender assignments that affect how the native speaker processes gender in their native language in real time (i.e., during speeded tasks). It is quite remarkable that the opposite trend was observed in adult learners in which being multilingual was associated with an advantage precisely with ungrammatical feminine nouns and speeded and auditory tasks. In other words, multilingual experience in native speakers appears to lead to interference and real-time (speeded) processing difficulties, while in adult learners, multilingual experience affords a noticeable advantage under increased processing demands, such as during speeded and auditory tasks, and with more difficult token types. Nonetheless, it is important to note that native speakers consistently performed at high levels of accuracy across all tasks and so perhaps proficiency in Spanish mitigates the nature and degree of the multilingual effect such that at lower proficiency levels, given that there is a linguistic deficit in the target language, any other linguistic resources available to the learner are *facilitative*—including knowledge of other languages—whereas once the target language grammar becomes more consolidated and complete, competing linguistic structures from other known languages may interfere with target language processing, particularly during increased processing constraints such as under a time pressure.

6.5.2 Multilingual effect: general typology

In order to investigate more precisely the nature of the positive effect of multilingualism in adult learners, multilingual learners were further subdivided according to whether or not they reported knowledge of an additional language beyond Spanish that also exhibits a grammatical gender system. This analysis allowed us to explore the question of whether or not the general effect of being multilingual may be independent from the effect of knowing another gendered language, thereby considering to what extent the observed multilingual learner advantage may be attributable to the transfer of a pre-existing grammatical gender system or to what extent the observed multilingual effect may be independent of language typology with respect to gender. Recall that multilingual native Spanish speakers were intentionally excluded from this analysis since their native language features a grammatical gender system and there is no objective way to tease apart the effect of a gendered L1 from the effect of a gendered L*n* (other additional language) in the present sample. Results indicated that the positive effect of multilingualism in adult language learners appears to be connected to the typology of their prior linguistic repertoire in that multilingual learners with grammatical gender “present” in their repertoire exhibited higher accuracy rates, particularly on auditory and speeded tasks, and higher accuracy on the more difficult ungrammatical feminine noun token type, than their multilingual learner peers whose prior linguistic repertoire does not feature a grammatical gender system. Therefore, the multilingual learner advantage observed in the present study could be explained to some extent by the transfer of a pre-existing gendered noun classification system. Nonetheless, a cross-language noun gender symmetry analysis is needed that compares accuracy scores for each multilingual learner on noun tokens both with the same and different gender assignment as compared to another reported gendered language. This type of noun gender symmetry analysis across the languages of multilingual learners would reveal if learners with multilingual linguistic knowledge are simply transferring the same gender assignment from an already known language into Spanish or if they demonstrate more general cognitive gains, which would be independent of matching gender assignment across languages.

6.5.3 Multilingual effect: gendered language subtype

The effect of gendered language subtype was also examined, that is, the presence of *binary* grammatical gender in the linguistic repertoire of the multilingual learner as compared to other manifestations of grammatical gender into more than two noun class categories. This analysis allowed us to explore further the nature of the observed multilingual learner advantage. Again, recall that multilingual native Spanish speakers were intentionally excluded from this analysis due to already having a binary gendered L1 (Spanish). Results indicated that multilingual learners with *binary* grammatical gender present in their prior linguistic repertoire did *not* perform better than their peers with other types of gendered noun class systems, and, in fact, the presence of binary grammatical gender appeared to be associated with lower average accuracy rates, particularly on speeded, auditory, and written tasks and furthermore did not appear to reduce group variation (SD) in performance. Finally, no effect of noun gender class system type was observed on the more difficult ungrammatical feminine noun token type—the token type believed to be most indicative of acquisition. Therefore, the observed typological multilingual learner advantage is limited to the presence of a grammatical gender system in the prior linguistic repertoire of the multilingual learner, but does not, however, appear to be associated with the particular subtype of gendered noun class system. Nonetheless, the present analysis is severely limited by unequal sampling of adult multilingual learners with (n=52) and without (n=4) binary grammatical gender present in their prior linguistic repertoire; among those multilingual learners with grammatical gender already present in their repertoire, the vast majority had binary grammatical gender present as well. Therefore, future research is needed from a more diverse sampling of multilingual learners who report proficiency in a variety of gendered languages of different subtypes.

6.5.4 Multilingual effect: number of additional languages

The multilingual effect was further examined by comparing multilingual participants (both native speakers and adult learners) according to the number of additional languages reported beyond Spanish and English. Given that a typological multilingual learner advantage was observed, this analysis aimed to demonstrate whether knowing two or three additional languages would afford greater advantage than knowing just one additional language. Results indicated that the number of additional languages known appears to have an effect on multilingual learner performance in which knowing two or more additional languages is more advantageous than knowing just one additional language, and this effect is particularly pronounced with the more difficult ungrammatical feminine noun token type (more indicative of acquisition) and with the more cognitively demanding auditory tasks (that tap into implicit linguistic competence). However, the effect of number of additional languages known on self-paced and written tasks is less evident. Furthermore, the observed *quantitative* multilingual learner advantage appears to be more binary in nature, observable when comparing multilingual learners reporting just one versus three additional languages, whereas much less of a differentiated effect is observed between multilingual learners reporting two and three additional languages. Therefore, we can conclude from the present multilingual learner dataset that there appears to be a *quantitative multilingual learner advantage* for knowing *two or more* additional languages on performance with grammatical gender in Spanish as a novel language. Moreover, given that this advantage is not evident on self-paced and written tasks, it may be more pertinent for learners' *implicit linguistic competence* employed during online language processing with speeded and auditory tasks. Nonetheless, this analysis is also limited by the unequal sampling of multilingual learners reporting one (n=44), two (n=24), and three (n=6) additional languages. Therefore, similarly to the typological multilingual analysis described above, future research is needed from a more diverse sampling of multilingual learners who report proficiency in three or more additional languages in order to confirm the preliminary findings uncovered here.

Regarding the quantitative effect of multilingualism in native Spanish speakers, the present sample was extremely limited as only one multilingual native speaker participant reported knowledge of two additional languages beyond Spanish and English and again only one native speaker reported knowledge of three additional languages, whereas the vast majority (n=10) of multilingual native speakers reported just one additional language. Nonetheless, the limited preliminary findings uncovered in the present study regarding the quantitative effect of multilingualism on native Spanish speaker performance seem to indicate an advantage for knowing three additional languages, particularly on the self-paced task type, contrary to the quantitative multilingual advantage observed in the adult learner group. Therefore, in native Spanish speakers, knowing three additional languages is more advantageous than knowing just one additional language, but this advantage appears to be within the domain of *explicit* linguistic knowledge that develops through extensive multilingual experience; however, on tasks that tap into implicit linguistic competence in multilingual native speakers (i.e., during online processing on speeded and auditory tasks), knowing more languages does not impact L1 performance. Nonetheless, we must emphasize that this analysis is extremely limited and inconclusive due to the unequal sampling of multilingual native speakers reporting one (n=10) as compared to two (n=1) and three (n=1) additional languages. Therefore, further research is needed to confirm these preliminary descriptive findings.

6.5.5 Multilingual effect: predictive modeling

In order to determine to what extent the observed multilingual effect may be significantly predictive of both learner and native speaker performance, linear regression models were constructed at a confidence interval of 95% for each participant group. A significant linear regression model was found for the adult learner group (n=90) in which the status of grammatical gender ("present" – "absent") in the individual's prior linguistic repertoire accounted for approximately 4% of the variation in learner accuracy scores. This was found to be the most significant and predictive model compared to other models that included the coefficients of 'multilingualism' and 'binary grammatical gender'. Therefore, it appears that accuracy scores with grammatical gender in Spanish as a late acquired nonnative language can be partially accounted for—to a very minimal extent (4%)—by the presence of a grammatical gender system in the prior linguistic repertoire of the multilingual adult learner. Evidently, the individual factors of proficiency, metalinguistic awareness, and motivation contribute notably more to learner performance than having a multilingual language background. Nonetheless, future research is needed to determine to what extent a multilingual learner background may interact with other learner characteristics which may be more developed in multilingual language learners, such as more positive attitudes, more developed metalinguistic and metacognitive awareness, and finally to what extent the proficiency level in multilinguals' other reported languages may modulate the observed multilingual advantage.

No significant linear regression model using multilingualism to predict performance was found for the native speaker group (n=25). Therefore, although the preliminary descriptive results seemed to indicate a slight multilingual advantage for native speakers on self-paced tasks in which they could tap into their declarative linguistic knowledge, this was not corroborated by inferential analysis. Although multilingualism is partially predictive of learner performance, multilingualism is *not* significantly predictive of performance with grammatical gender in Spanish as a native language.

6.5.6: Multilingual effect: possible explanations

We will now consider possible explanations for the observed typological and quantitative multilingual advantage that is slightly predictive of late learner performance. Recall that multilingual speakers have an empirically demonstrated advantage when it comes to the task of learning a third (L3) or additional (L*n*) language in lexical (e.g., Kaushanskaya & Marian, 2009), phonological (e.g., Tremblay & Sabourin, 2012), phonetic (e.g., Antoniou et al., 2015), and syntactic (e.g., Klein, 1995) domains. A typological and quantitative multilingual advantage was also detected in the present study, thereby adding to this growing body of evidence for a multilingual late learner advantage for the learning of a specific morphosyntactic structure. In terms of how this multilingual advantage could be explained, previous researchers have posited that metalinguistic awareness and linguistic repertoire (e.g., Cenoz, 2013), affective factors (Dewaele et al., 2008), as well as general changes to the cognitive-linguistic system (Hirosh & Degani, 2018) could explain to some extent why being multilingual affords an advantage for subsequent language learning. Kemp (2007) argues that multilingual learners may develop a higher level of metalinguistic awareness on the basis of their previous language learning experience and their knowledge of more than one linguistic system. The idea here is that multilingual learners are able to think about language in a more abstract way and regard it as an object of analysis more so than their peers with only (largely) implicit linguistic competence in their native language (Cenoz, 2013). In other words, one possible explanation for the multilingual learner advantage is that multilingual speakers possess explicit/declarative knowledge about language to a larger extent than their monolingual and bilingual peers. Nonetheless, our findings do *not* corroborate this potential explanation as no significant correlation was detected between multilingualism and metalinguistic awareness nor attitudes, indicating that multilingual learners are not more likely to have more metalinguistic awareness in the language nor have more positive attitudes about the target language community. Therefore, these individual factors cannot explain the observed multilingual advantage. Moreover, a larger multilingual advantage was detected on online/speeded tasks that tap into *implicit* linguistic competence. Therefore, we find that being multilingual is associated with more developed implicit linguistic competence evident during speeded language processing, contrary to the possible explanations of Cenoz (2013) and Kemp (2007). Our findings more closely align with Mady (2017) who found that immigrant multilingual students in a French immersion program in Canada outperformed their Canadian-born bilingual peers on French proficiency tests although no significant differences were detected between these two groups with regards to metalinguistic knowledge nor strategy use, therefore suggesting that the multilingual advantage observed could not be attributable to difference in explicit language knowledge. Evidently, we must look elsewhere to elucidate the nature of the multilingual effect.

Cenoz (2013) also argues that the multilingual advantage in additional language learning can be largely attributed to multilingual learners' larger linguistic and intercultural repertoires which they can employ to their advantage in L*n* learning. The present findings seem to collaborate Cenoz's explanation for the role of the multilingual learner's linguistic repertoire as our findings suggest a typological advantage for those multilingual learners with a grammatical gender system already present in their prior linguistic repertoire. Nonetheless, this advantage only holds for the general presence of grammatical gender in their linguistic repertoire, independently of the particular gendered noun class system subtype. Cenoz (2013) also argues that this advantage could be linked to more extensive experience with language learning as a specific set of skills such that multilinguals tend to have more developed and effective language learning strategies that they have refined over time. Our findings are not able to respond to this particular factor of the multilingual learner background as we did not collect information on language learning skills specifically. Therefore, skills and strategy use as a potential explanatory factor for the observed multilingual learner advantage requires further research.

Some researchers have argued that the effect of linguistic repertoire on multilingual language learning is really a question of *language distance* or *typology*; closely related languages are more useful for multilingual learners when learning an additional language (Ringbom, 2007; Jarvis & Pavlenko, 2008). The degree of typological overlap between the languages of the multilingual speaker may play an important role in the pattern and quantity of transfer when learning a novel language (see Rothman, 2013, 2014; Rothman, Cabrelli Amaro, & de Bot, 2013 for a discussion on transfer source related to typology in third language acquisition) and the effect of language similarity may surface when the structure in question is particularly difficult to acquire (Hirosh & Degani, 2008), as is the case with grammatical gender as a problematic structure for late acquisition. Our findings seem to corroborate this explanation of typological similarity. In the present sample of multilingual learners, the positive effect of multilingualism appears to be typological in nature in which multilingual learners with grammatical gender (of any subtype) already present in their prior linguistic repertoire outperformed their multilingual peers who did not have grammatical gender already present. Nonetheless, linear regression modeling revealed that having grammatical gender present in one's linguistic repertoire is only *minimally* predictive of performance in late learners such that only 4% of the variation in performance can be explained by this factor alone. Therefore, our descriptive findings corroborate previous research that establishes a multilingual advantage, yet multilingualism as a predictor of late learner performance appears to only be minimally predictive of performance. Nonetheless, further research is needed to explore in what ways not accounted for in the present analysis the multilingual learner group may be different from their bilingual/monolingual peers. In other words, what other individual characteristics, such as enhanced strategy use, tend to occur more frequently and to a larger extent in multilingual learners and therefore may help to further elucidate the nature of the multilingual advantage detected here and in previous work.

Hirosh and Degani (2018) uncovered two main subcomponents of the multilingual advantage in their systematic review of thirty-three empirical studies, including direct and indirect influence. *Direct* influence entails direct transfer of prior knowledge and skills while *indirect* influence involves more general changes to the cognitive-linguistic system, including linguistic and non-linguistic executive function leading to the enhanced cognitive abilities associated with multilingualism. Furthermore, Hirosh and Degani (2018) assert that prior language learning context, that is, instructed/formal as compared to naturalistic/informal, appears to influence the balance between the direct and indirect effects of multilingualism on novel language learning such that a formal language learning context (i.e., classroom) tends to favor more direct multilingual effects, including the direct transfer of grammatical knowledge and learning strategies, whereas individuals exposed informally to additional languages tend to rely more on the indirect effects of multilingualism, including enhanced inhibition and attention control, ambiguity processing, verbal memory, and a more developed lexical-semantic network. Nonetheless, our data cannot directly elucidate this distinction since we did not collect information on the learning contexts of multilingual participants in their additional reported languages. Therefore, future work is needed in which the context of acquisition of additional languages is considered for the interpretation of the multilingual learner advantage. However, it is important to note, as discussed previously, that the multilingual learner advantage detected in the present study was both typological in nature (i.e., grammatical gender in the prior linguistic repertoire) and associated with enhanced implicit linguistic competence (i.e., enhanced advantage on speeded and auditory tasks). Furthermore, as Hirosh and Degani (2018) note, research is still lacking in the domain of monolingual-multilingual differences in learning the grammar of a novel language. Therefore, this study makes a novel contribution to the growing body of research on multilingual language acquisition by demonstrating a multilingual learner advantage for performance with a particularly problematic morphosyntactic structure that is related both to the typology of other languages known (direct transfer of grammatical structure) as well as associated with enhanced implicit linguistic competence (indirect changes to the cognitive-linguistic system).

In sum, the observed typological and quantitative multilingual learner advantage associated with the presence of a grammatical gender system in the prior linguistic repertoire and the modulating effect of the number of additional languages, in addition to the finding that this multilingual advantage is minimally predictive of performance in late learners and does not predict native speaker performance, can be partially explained by changes to the cognitive-linguistic system through the direct and indirect influence of multilingualism (Hirosh & Degani, 2018), the role of typology and linguistic repertoire (Cenoz, 2013), and by the concept of language distance (Ringbom, 2007; Jarvis & Pavlenko, 2008).

We will now examine the combined effect of all independent explanatory variables through comprehensive linear regression modeling to explore how performance with grammatical gender in Spanish as both a native and late acquired language can be predicted with these factors.

6.6 Comprehensive predictive modeling

Although each research question up until now has addressed separately the effect on performance of the linguistic, task, and individual factors, including multilingualism, to varying results, it is relevant to reexamine the predictive power of all these factors taken together. In other words, to what extent can we take these factors and use them to predict performance with a particularly problematic linguistic structure? To what extent can we model how both adult/late learners and native speakers may be variably impacted by these distinct factors? What would this model reveal about the nature of both native and nonnative late-acquired language processing and performance? In order to address these concerns, multiple linear regression models were constructed at a confidence interval of 95% for each proficiency group—beginner, intermediate, advanced—of adult learners as well as all learners together and for all native speakers taken together using as model coefficients the linguistic factors (ungrammatical, feminine, and low frequency token scores), task factors (speeded and auditory task scores), and individual factors (Spanish proficiency, metalinguistic awareness, motivation, multilingualism, and prior linguistic repertoire) that had previously been identified as significantly predictive of performance in the present dataset when these factor types were analyzed separately. The linear regression model output for all adult learners combined is presented in Table 32, Table 33 presents the model output for beginner proficiency learners only, Table 34 presents the model output for intermediate proficiency learners only, and Table 35 presents the model output for advanced proficiency learners only. Finally, Table 36 presents the model output for native speakers only.

For late/adult learners considered *all together*, a significant multiple linear regression model was found that accounts for approximately 98% of the variation in average accuracy scores (adjusted $R^2 = 0.984$, $F(10, 79) = 536$, $p <.001$). The model output is presented below in Table 32. The following model coefficients were included as predictors in the final model: average ungrammatical token score, average feminine token score, low-frequency noun score, average speeded task score, average auditory task score, total tested Spanish proficiency score, total metalinguistic awareness score, total motivational score, multilingualism, and linguistic repertoire (i.e., the presence of grammatical gender). Of these final model coefficients, the following factors were found to be individually predictive of performance in descending order of effect size: low-frequency noun score ($F = 47.75$, $p < .001$), speeded score ($F = 19.35$, $p < .001$), total metalinguistic awareness score ($F = 16.54$, $p < .001$), and average ungrammatical token score ($F = 6.87$, $p = .011$)[4]. Therefore, when we consider all independent explanatory variables—including linguistic variables, task factors, and individual differences—in late/adult learners as a whole, we can conclude from the present dataset that learners' performance can be most predicted by *learner-external* variables, including the frequency of the task stimuli (i.e., relative target noun frequency in this case), the token grammaticality (i.e., ungrammatical token score), to a more minimal extent by the markedness of the target noun (i.e., feminine nouns appear to be marked or non-default), and the linguistic processing type conditioned by the time constraint nature of the task (i.e., speeded tasks), as well as by the *learner-internal* variable of how explicitly aware one is of the target structure in question (i.e., metalinguistic awareness). Nonetheless, it is just as interesting to note that while contributing to the predictive power of the model overall, tested Spanish proficiency, stimuli modality (i.e., auditory), motivation, being multilingual, and one's prior linguistic repertoire are *not* independently predictive of performance in late/adult learners as a whole.

[4] Note that average feminine noun token score almost reached significance ($F = 3.62$, $p = .061$) for the comprehensive linear regression model of nonnative learner performance.

Table 32. Output of a linear regression model (CI 95%) examining the combined effect of all factors previously identified as significant on average task scores for **all adult learners combined**. Individually significant model coefficient predictors are highlighted in grey.

				Overall Model Test	
Model	***R***	***R^2***	**Adjusted *R^2***	***F***	***p***
1	0.993	0.985	0.984	536	< .001

Omnibus ANOVA Test

	Sum of Squares	df	Mean Square	*F*	*p*
AVG UNGRAMMATICAL SCORE	0.00221	1	0.00221	6.870	0.011
AVG FEM SCORE	0.00116	1	0.00116	3.620	0.061
NOUN FREQ. SCORE: LOW	0.01535	1	0.01535	47.750	< .001
SPEEDED SCORE	0.00622	1	0.00622	19.347	< .001
AUDITORY SCORE	3.50e-4	1	3.50e-4	1.088	0.300
SPAN PROF: TOTAL SCORE	4.86e-4	1	4.86e-4	1.511	0.223
METALING AWARE: TOTAL	0.00532	1	0.00532	16.541	< .001
MOTIV: TOTAL SCORE	4.75e-5	1	4.75e-5	0.148	0.702
MULTILINGUALISM	2.41e-4	1	2.41e-4	0.751	0.389
GRAMM GEN PRESENT	2.53e-4	1	2.53e-4	0.787	0.378
Residuals	0.02540	79	3.22e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	0.06730	0.02423	0.01907	0.11553	2.778	0.007
AVG UNGRAMMATICAL SCORE	-0.08524	0.03252	-0.14998	-0.02051	-2.621	0.011
AVG FEM SCORE	0.07964	0.04186	-0.00368	0.16296	1.903	0.061
NOUN FREQ. SCORE: LOW	0.45088	0.06525	0.32100	0.58075	6.910	< .001
SPEEDED SCORE	0.56005	0.12733	0.30661	0.81349	4.398	< .001
AUDITORY SCORE	-0.08824	0.08461	-0.25666	0.08018	-1.043	0.300
SPAN PROF: TOTAL SCORE	-0.01888	0.01536	-0.04946	0.01169	-1.229	0.223
METALING AWARE: TOTAL	0.03006	0.00739	0.01535	0.04477	4.067	< .001
MOTIV: TOTAL SCORE	0.00136	0.00354	-0.00568	0.00840	0.384	0.702

MULTILINGUALISM:						
No – Yes	0.00560	0.00647	-0.00727	0.01847	0.866	0.389
GRAMM GEN PRESENT:						
No – Yes	-0.00466	0.00526	-0.01512	0.00580	-0.887	0.378

A significant multiple linear regression model was also found for adult *beginner* level proficiency learners that accounts for approximately 97% of the variation in average accuracy scores (adjusted $R^2 = 0.971$, $F(10, 51) = 203$, $p < .001$). The model output is presented below in Table 33. As with all the learner proficiency groups, the following model coefficients were included as predictors in the final model: average ungrammatical token score, average feminine token score, low-frequency noun score, average speeded task score, average auditory task score, total tested Spanish proficiency score, total metalinguistic awareness score, total motivational score, multilingualism, and linguistic repertoire (i.e., the presence of grammatical gender). Of these final model coefficients, the following factors were found to be individually predictive of performance in descending order of effect size: low-frequency noun score ($F = 37.80$, $p < .001$), average speeded task score ($F = 16.89$, $p < .001$), total metalinguistic awareness score ($F = 16.47$, $p < .001$), and average ungrammatical token score ($F = 3.79$, $p = .05$). Therefore, when we consider all independent explanatory variables—including linguistic variables, task factors, and individual differences—in late/adult learners at the beginner proficiency level, we can conclude from the present dataset that beginner learners' performance can be most predicted by the learner-external variables, including the frequency of the task stimuli (i.e., relative target noun frequency in this case), the token grammaticality (i.e., ungrammatical token score), and the linguistic processing type conditioned by the time constraint nature of the task (i.e., speeded tasks), as well as by the learner-internal variable of how explicitly aware one is of the target structure in question (i.e., metalinguistic awareness). Nonetheless, it is just as interesting to note that while contributing to the predictive power of the model overall, average feminine token score, average auditory task score, total tested Spanish proficiency, total motivational score, multilingualism, and linguistic repertoire are *not* independently predictive of performance in learners at the beginner proficiency level.

Table 33. Output of a linear regression model (CI 95%) examining the combined effect of all factors previously identified as significant on average task scores for **beginner proficiency learners only (n=62)**. Individually significant model coefficient predictors are highlighted in grey.

				Overall Model Test	
Model	***R***	***R²***	**Adjusted *R²***	***F***	***p***
1	0.988	0.975	0.971	203	< .001

Omnibus ANOVA Test

	Sum of Squares	df	Mean Square	*F*	*p*
AVG UNGRAMMATICAL SCORE	0.00130	1	0.00130	3.7917	0.050
AVG FEM SCORE	1.33e-4	1	1.33e-4	0.3886	0.536
NOUN FREQ. SCORE: LOW	0.01293	1	0.01293	37.7994	< .001
SPEEDED SCORE	0.00578	1	0.00578	16.8873	< .001
AUDITORY SCORE	2.77e-4	1	2.77e-4	0.8090	0.373
SPAN PROF: TOTAL SCORE	4.24e-4	1	4.24e-4	1.2407	0.271
METALING AWARE: TOTAL	0.00563	1	0.00563	16.4715	< .001
MOTIV: TOTAL SCORE	7.01e-6	1	7.01e-6	0.0205	0.887
MULTILINGUALISM	9.40e-5	1	9.40e-5	0.2748	0.602
GRAMM GEN PRESENT	2.07e-5	1	2.07e-5	0.0605	0.807
Residuals	0.01745	51	3.42e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	0.06534	0.02848	0.00817	0.12251	2.295	0.026
AVG UNGRAMMATICAL SCORE	-0.07286	0.03742	-0.14799	0.00226	-1.947	0.053
AVG FEM SCORE	0.03340	0.05358	-0.07417	0.14098	0.623	0.536
NOUN FREQ. SCORE: LOW	0.44559	0.07248	0.30009	0.59109	6.148	< .001
SPEEDED SCORE	0.60054	0.14614	0.30715	0.89392	4.109	< .001
AUDITORY SCORE	-0.08719	0.09693	-0.28179	0.10741	-0.899	0.373
SPAN PROF: TOTAL SCORE	-0.03261	0.02927	-0.09137	0.02616	-1.114	0.271
METALING AWARE: TOTAL	0.03411	0.00840	0.01724	0.05098	4.059	< .001
MOTIV: TOTAL SCORE	-5.90e−4	0.00412	-0.00887	0.00769	-0.143	0.887

MULTILINGUALISM:						
No – Yes	0.00403	0.00768	-0.01139	0.01945	0.524	0.602
GRAMM GEN PRESENT:						
No – Yes	-0.00153	0.00623	-0.01405	0.01098	-0.246	0.807

A significant multiple linear regression model was <u>not</u> found for *intermediate* proficiency adult learners (adjusted R^2 = 0.794, $F(10, 3)$ = 6.00, p = .084). The model output is presented below in Table 34. As with all the learner proficiency groups, the following model coefficients were included as predictors in the final model: average ungrammatical token score, average feminine token score, low-frequency noun score, average speeded task score, average auditory task score, total tested Spanish proficiency score, total metalinguistic awareness score, total motivational score, multilingualism, and linguistic repertoire (i.e., the presence of grammatical gender). Of these final model coefficients, *none* of the factors were found to be individually predictive of performance. Therefore, when we consider all independent explanatory variables—including linguistic variables, task factors, and individual differences—in late/adult learners at the intermediate proficiency level, we can conclude from the present dataset that intermediate learners' performance is *not* able to be significantly predicted by any of the individual factors that were previously identified as significant when all adult learners are considered together.

Table 34. Output of a linear regression model (CI 95%) examining the combined effect of all factors previously identified as significant on average task scores for **intermediate proficiency learners only (n=14)**.

				Overall Model Test	
Model	***R***	***R²***	**Adjusted *R²***	***F***	***p***
1	0.976	0.952	0.794	6.00	0.084

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
AVG UNGRAMMATICAL SCORE	2.49e-4	1	2.49e-4	0.4116	0.567
AVG FEM SCORE	4.35e-4	1	4.35e-4	0.7194	0.459
NOUN FREQ. SCORE: LOW	0.00200	1	0.00200	3.3104	0.166
SPEEDED SCORE	1.41e-4	1	1.41e-4	0.2330	0.662
AUDITORY SCORE	2.21e-4	1	2.21e-4	0.3654	0.588

SPAN PROF: TOTAL SCORE	4.42e-4	1	4.42e-4	0.7318	0.455
METALING AWARE: TOTAL	7.15e-5	1	7.15e-5	0.1183	0.754
MOTIV: TOTAL SCORE	9.11e-4	1	9.11e-4	1.5077	0.307
MULTILINGUALISM	1.34e-4	1	1.34e-4	0.2220	0.670
GRAMM GEN PRESENT	1.15e-5	1	1.15e-5	0.0190	0.899
Residuals	0.00181	3	6.04e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	0.32397	0.2728	-0.5443	1.192	1.187	0.321
AVG UNGRAMMATICAL SCORE	-0.13382	0.2086	-0.7976	0.530	-0.642	0.567
AVG FEM SCORE	0.21590	0.2546	-0.5942	1.026	0.848	0.459
NOUN FREQ. SCORE: LOW	0.63703	0.3501	-0.4772	1.751	1.819	0.166
SPEEDED SCORE	0.48388	1.0024	-2.7061	3.674	0.483	0.662
AUDITORY SCORE	-0.45314	0.7496	-2.8387	1.932	-0.605	0.588
SPAN PROF: TOTAL SCORE	-0.19275	0.2253	-0.9098	0.524	-0.855	0.455
METALING AWARE: TOTAL	0.02661	0.0774	-0.2196	0.273	0.344	0.754
MOTIV: TOTAL SCORE	0.03404	0.0277	-0.0542	0.122	1.228	0.307
MULTILINGUALISM:						
No – Yes	-0.03742	0.0794	-0.2902	0.215	-0.471	0.670
GRAMM GEN PRESENT:						
No – Yes	0.00514	0.0373	-0.1135	0.124	0.138	0.899

Nonetheless, a significant multiple linear regression model was found for *advanced* proficiency adult learners that accounts for approximately 94% of the variation in average accuracy scores (adjusted $R^2 = 0.940$, $F(10, 3) = 21.4$, $p = .014$). The model output is presented below in Table 35. As with all the learner proficiency groups, the following model coefficients were included as predictors in the final model: average ungrammatical token score, average feminine token score, low-frequency noun score, average speeded task score, average auditory task score, total tested Spanish proficiency score, total metalinguistic awareness score, total motivational score, multilingualism, and linguistic repertoire (i.e., the presence of grammatical gender). Although the model as a whole was found to be a significant predictor of performance, *none* of the individual model coefficients were found to be *individually* predictive of performance. Therefore, when we consider all independent explanatory variables—including linguistic variables, task factors, and individual differences—in late/adult learners at the advanced proficiency level, we can conclude from the present dataset that advanced learners' performance can be significantly predicted by a variety of *learner-external* factors, such as frequency of the task stimuli, token grammaticality, and the linguistic processing type conditioned by the time constraint nature of the task, as well as by *learner-internal* factors, such as how explicitly aware one is of the target structure in question (i.e., metalinguistic awareness), one's current target proficiency level, and how many languages one already knows and if they already have a grammatical gender system present in their mental grammar repertoire, although the exact predictive power of each factor is non-significant with the present sample. Perhaps with more advanced proficiency learner data the factors identified in the present predictive modeling would be found to be individually significant. For now, we can only conclude that they have some influence, but are crucially not independently predictive of nonnative linguistic performance in Spanish at the advanced proficiency level.

Table 35. Output of a linear regression model (CI 95%) examining the combined effect of all factors previously identified as significant on average task scores for **advanced proficiency learners only (n=14)**. Note that the *F*-values of the factors average speeded task score, total Spanish proficiency, total motivational score, and multilingualism seem to indicate that they contribute to the predictive power of the model overall, although they are not *individually* predictive of performance, according to their respective *p*-values.

				Overall Model Test	
Model	***R***	***R²***	**Adjusted *R²***	***F***	***p***
1	0.993	0.986	0.940	21.4	0.014

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
AVG UNGRAMMATICAL SCORE	4.52e-5	1	4.52e-5	0.347	0.597
AVG FEM SCORE	4.66e-5	1	4.66e-5	0.357	0.592
NOUN FREQ. SCORE: LOW	4.31e-5	1	4.31e-5	0.330	0.606
SPEEDED SCORE	2.24e-4	1	2.24e-4	1.717	0.281
AUDITORY SCORE	1.52e-5	1	1.52e-5	0.116	0.756
SPAN PROF: TOTAL SCORE	3.10e-4	1	3.10e-4	2.377	0.221

METALING AWARE: TOTAL	2.20e-5	1	2.20e-5	0.168	0.709
MOTIV: TOTAL SCORE	1.33e-4	1	1.33e-4	1.018	0.387
MULTILINGUALISM	1.31e-4	1	1.31e-4	1.006	0.390
GRAMM GEN PRESENT	1.06e-4	1	1.06e-4	0.813	0.434
Residuals	3.91e-4	3	1.30e-4		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	-0.3798	0.3169	-1.3884	0.6288	-1.198	0.317
AVG UNGRAMMATICAL SCORE	0.2771	0.4704	-1.2199	1.7742	0.589	0.597
AVG FEM SCORE	-0.1894	0.3170	-1.1981	0.8192	-0.598	0.592
NOUN FREQ. SCORE: LOW	-0.3177	0.5527	-2.0766	1.4412	-0.575	0.606
SPEEDED SCORE	1.6977	1.2955	-2.4252	5.8206	1.310	0.281
AUDITORY SCORE	-0.2694	0.7899	-2.7832	2.2445	-0.341	0.756
SPAN PROF: TOTAL SCORE	0.1530	0.0992	-0.1628	0.4687	1.542	0.221
METALING AWARE: TOTAL	0.0148	0.0360	-0.0999	0.1295	0.410	0.709
MOTIV: TOTAL SCORE	0.0238	0.0236	-0.0513	0.0990	1.009	0.387
MULTILINGUALISM:						
No – Yes	-0.0568	0.0566	-0.2370	0.1234	-1.003	0.390
GRAMM GEN PRESENT:						
No – Yes	0.0414	0.0460	-0.1048	0.1877	0.902	0.434

Finally, a significant multiple linear regression model was also found for the adult *native Spanish speakers* that accounts for approximately 95% of the variation in average accuracy scores (adjusted $R^2 = 0.953$, $F(5, 19) = 98.3$, p <.001). The model output is presented below in Table 36. The following model coefficients were included as predictors in the final model to predict native speaker performance: noun frequency-average score correlation, Spanish use score, average speeded task score, total metalinguistic awareness score, and multilingualism. Of these final model coefficients, the following factors were found to be individually predictive of performance in descending order of effect size: average speeded task score ($F = 367.37$, $p < .001$), noun frequency-average score correlation ($F = 10.11$, $p = .005$), and Spanish use score ($F = 6.49$, $p =$

.020). Therefore, when we consider all independent explanatory variables—including linguistic variables, task factors, and individual differences—in adult native speakers, we can conclude from the present dataset that native speaker performance can be predicted to a large extent by *speaker-external* factors, such as the linguistic processing type conditioned by the time constraint nature of the task and the frequency of the task stimuli, as well as by the *speaker-internal* factor of how often the native speaker chooses to use their native language across different contexts in an average week. Nonetheless, it is just as interesting to note that while contributing to the predictive power of the model overall, total metalinguistic awareness scores and knowing additional languages beyond Spanish and English (i.e., multilingual), are not *independently* predictive of performance with grammatical gender in Spanish as a native language.

Table 36. Output of a linear regression model (CI 95%) examining the combined effect of all factors previously identified as significant on average task scores for **native speakers only (n=25)**. Individually significant model coefficient predictors are highlighted in grey.

				Overall Model Test	
Model	***R***	***R^2***	**Adjusted *R^2***	***F***	***p***
1	0.981	0.963	0.953	98.3	< .001

Omnibus ANOVA Test

	Sum of Squares	**df**	**Mean Square**	***F***	***p***
NOUN FREQ.: SCORE CORRELATION	2.55e-4	1	2.55e-4	10.11	0.005
SPAN USE SCORE	1.64e-4	1	1.64e-4	6.49	0.020
SPEEDED SCORE	0.00925	1	0.00925	367.37	< .001
METALING AWARE: OVERT	3.43e-5	1	3.43e-5	1.36	0.258
MULTILINGUALISM	6.61e-5	1	6.61e-5	2.62	0.122
Residuals	4.79e-4	19	2.52e-5		

Note. Type 3 sum of squares

Model Coefficients - AVG TASK SCORE

			95% Confidence Interval			
Predictor	**Estimate**	**SE**	**Lower**	**Upper**	***t***	***p***
Intercept [a]	0.25434	0.03752	0.17581	0.33286	6.78	< .001
NOUN FREQ.: SCORE CORRELATION	-0.05118	0.01610	-0.08487	-0.01749	-3.18	0.005
SPAN USE SCORE	-0.01534	0.00602	-0.02795	-0.00274	-2.55	0.020
SPEEDED SCORE	0.75164	0.03922	0.66956	0.83372	19.17	< .001

METALING AWARE: OVERT	0.00270	0.00232	-0.00215	0.00755	1.17	0.258
MULTILINGUALISM:						
No – Yes	-0.00401	0.00247	-0.00919	0.00117	-1.62	0.122

Table 37 summarizes the *learner/speaker-external* and *learner/speaker-internal* factors that have been identified as *individually* significant predictors of performance with grammatical gender in Spanish as both a native and nonnative/late-acquired language and delineates how these significant predictors vary according to learner proficiency level. It is important to note that none of the individual model coefficients for the intermediate proficiency learner group were significantly predictive of performance nor was their combined predictive value significant for the model overall (see *Table 34*). In contrast, for the advanced proficiency learners, although none of the model coefficients were individually predictive of performance, their combined predictive value was indeed found to be significant (see *Table 35*).

Table 37. Summary of factors found to be individually significant predictors of performance with Spanish grammatical gender according to multiple linear regression modeling in the following participant groups: native speakers, all learners combined, beginner proficiency learners, intermediate proficiency learners, and advanced proficiency learners. Individually significant performance predictors for each group are highlighted in grey.

Significant Individual Performance Predictor:	**Native Speakers (n=25)**	**All Learners (n=90)**	**Beginner Proficiency (n=62)**	**Intermediate Proficiency (n=14)**	**Advanced Proficiency (n=14)**
EXTERNAL FACTORS:					
token grammaticality	**NO**	**√ YES**	**√ YES**	**NO**	**NO**
noun frequency	**√ YES**	**√ YES**	**√ YES**	**NO**	**NO**
task processing type conditioned	**√ YES**	**√ YES**	**√ YES**	**NO**	**NO**
INTERNAL FACTORS:					
metalinguistic awareness	**NO**	**√ YES**	**√ YES**	**NO**	**NO**
target language use	**√ YES**	**NO**	**NO**	**NO**	**NO**

When considered together, these models tell us that both native speakers and adult learners of Spanish are affected by different external and internal factors that can be used either individually or together to predict performance. More specifically, both native speaker and adult learner performance can be significantly and independently predicted by the relative frequency of the target nouns and the task processing type that is conditioned. However, only adult learner performance can also be predicted by the grammaticality of the target token as well as by learners' explicit awareness of grammatical gender as the target structure (i.e., metalinguistic awareness), and only native speaker performance can be partially predicted by average target language use. Furthermore, in adult learners, proficiency level appears to modulate the relative predictive value that these factors have such that at intermediate levels of proficiency, none of the identified factors are predictive of performance—not individually nor when considered together—whereas for advanced proficiency adult learners, although the identified factors are not individually predictive of performance, when considered together they can significantly predict performance.

With regards to token grammaticality, it appears that in the present dataset, 'grammatical' was treated as default response in adult learners likely to compensate for a linguistic deficit whereas token grammaticality is not predictive of native speaker performance as native speakers do not demonstrate a default response as no compensation strategy is needed likely because the adult native speakers have already fully acquired their native language, including grammatical gender. Nonetheless, if we tested native speaker performance with unfamiliar or nonce nouns—nouns that they do not currently have in their lexical inventory—perhaps we would find a similar default processing strategy, but this would probably be somewhat related to morphological cues present in the unknown noun token (i.e., morphological transparency for gender).

Regarding the predictive power of noun frequency, language processing is based on knowledge about relative frequencies that allows for probabilistic judgements to occur largely without our awareness during implicit learning and that are subsequently exploited during implicit processing of language (N.C. Ellis, 2002). Therefore, the fact that noun frequency is an individually significant predictor of performance in both adult learners and native speakers alike tells us something about the nature of the grammatical knowledge/competence that is being exploited in order to produce an average accuracy score with grammatical gender. Of course, the average accuracy score takes into account performance on all tasks—ones that condition both implicit and explicit language processing—and, therefore, if we used 'average speeded score' as the dependent response variable and compared the predictive value of the model to a model that predicts ‘average self-paced score’ as the dependent response variable, we might find that the model including the coefficient of ‘noun frequency’ to be more predictive of performance with average speeded scores than with average self-paced scores, but it is also likely that this effect would be similar for both native speakers and nonnative learners, although recall that the direction of the task processing modality effect is reversed in native speakers (i.e., they perform better on speeded tasks than on self-paced tasks).

With regards to task processing type as a predictor of performance, we know that certain task design features, namely the presence/absence of a time constraint, differentially conditions the deployment of distinct knowledge and memory systems; therefore, it is likely that introducing a time constraint into the task design conditioned both learners and native speakers alike to access their implicit linguistic competence to varying degrees of success, thus serving as a significant predictor of performance as those language users with more developed implicit linguistic competence (i.e., advanced learners and native speakers) were able to, to a larger extent, accurately respond to the task prompts in a speeded manner, whereas beginner learners also accessed their implicit linguistic competence that is still in a state of flux and development (i.e., correspondingly higher intra-speaker variance) and, therefore, their respective speeded performance reflected this linguistic deficit relative to their advanced proficiency and native speaker peers. Recall that according to the descriptive data per participant group, native speakers and advanced proficiency adult learners performed notably better on speeded tasks while intermediate and beginner learners performed notably better on self-paced tasks. Therefore, when we consider the predictive value of task processing type along with respective accuracy scores on speeded as compared to self-paced tasks across participant groups, we can conclude that native speaker and advanced learner performance is supported to a larger extent by *implicit linguistic competence* in procedural memory while adult learner performance at lower levels of proficiency is supported to a larger extent by *explicit linguistic knowledge* in declarative memory. In other words, the linguistic performance of native and nonnative speakers alike is modulated by the language processing type that is conditioned by the task design such that the time constraint factor predicts significantly *lower* accuracy in beginner and intermediate adult learners and significantly *higher* accuracy in native speakers and advanced proficiency learners, revealing the relative balance between the distinct knowledge and memory systems employed by both native speakers and adult instructed language learners. Nonetheless, as Ullman (2015; 2012; 2005) argues, both declarative and procedural long-term memory are exploited in both first and second language acquisition but differ in their degree of use at any point in time such that early language acquisition depends more on the procedural long-term memory whereas subsequent late language acquisition tends to depend more on declarative long-term memory. Therefore, all language users use both implicit linguistic competence in procedural memory and explicit linguistic knowledge in declarative memory, although the relative balance between these knowledge and memory systems differs and evolves through time. In sum, the predictive modeling in the present dataset suggests that *beginner* and *intermediate* proficiency learners depend to a larger extent on *explicit* language knowledge stored in *declarative* memory and *native speakers* as well as *advanced* learners depend to a larger extent on *implicit* linguistic competence stored in *procedural* memory to support their respective linguistic performance.

In line with the implicit-explicit divide detected in participant performance, metalinguistic awareness was also found to be a significant predictive factor. It makes sense that one's level of explicit awareness of the target structure (i.e., metalinguistic awareness) is predictive of performance in adult learners, but crucially *not* significantly predictive of performance in native speakers, since late learners are posited to rely more extensively on metalinguistic knowledge as a compensation strategy for gaps in their implicit linguistic competence. This indeed seems to be corroborated by the present linear regression predictive modeling as metalinguistic awareness appears to be predictive of performance in *beginner* proficiency learners only, and this effect subsides at intermediate and advanced levels of proficiency when explicit awareness is less needed—and can even impede linguistic performance, particularly during speeded/online processing, as observed in the native speaker and advanced learner groups.

Regarding the modulating effect of proficiency, when learners of all proficiency levels are considered together, more individual significant predictors of performance can be identified (i.e., token grammaticality, noun frequency, task processing type, metalinguistic awareness), but as proficiency increases, these predictors are no longer significant. There is evidently something else going on here. Most obviously, this lack of significant predictive factors in intermediate and advanced proficiency learners is linked to the relatively small sample size of these proficiency groups, that is, only 14 learner participants had tested Spanish proficiency scores that placed them in the 'intermediate group' and, likewise, only 14 other participants had tested Spanish proficiency scores that placed them in the 'advanced group'. For predictive modeling, both of these groups were individually compared to the 'beginner group', in which 62 participants were placed according to their tested Spanish proficiency scores. Therefore, there might be too much variation relative to the sample size impeding a significant *p*-value for inferential testing in the intermediate and advanced learner groups. Although notable inter-learner variation (SD) was observed in the beginner proficiency group, the fact that this group was over four times larger than the other two proficiency groups likely helped in producing significant outcomes for inferential testing. Nonetheless, there still could be other confounding factors not yet uncovered in the present sample that are potentially predictive of performance at more advanced levels of proficiency such as context of learning (solely instructed or instructed/naturalistic mix), amount and nature of language use outside of the classroom (e.g., study abroad, family, friends, work), vocabulary size, or even the different subdomains of motivation and attitudes as these were not directly entered into predictive modeling (only overall motivational and attitude scores were considered).

Therefore, from this modeling with the present sample of intermediate and advanced adult learners of Spanish, we cannot draw meaningful conclusions about how learners' sensitivities to these predictive factors may evolve as overall proficiency in the language increases. To address this deficit, further research is needed with a larger intermediate and advanced learner sample size and the different subdomains of each of the self-reported individual factors could also be entered into predictive modeling to determine whether significant predictors may be uncovered for intermediate and advanced learners that were not detected in beginner proficiency learners of Spanish.

Furthermore, it is not clear why target language use would be a significant individual predictor of performance in native speakers but not significantly predict the performance of adult learners. It seems reasonable to assume that if one uses the target language more, one would have more opportunities to process comprehensible input to add to and refine their developing linguistic system or interlanguage and thereby enhance their performance. Likewise, it seems counterintuitive that native speakers would need this continued language use to reinforce their performance with a grammatical structure that is posited to be acquired at around age 3. Therefore, from the current predictive modeling from the present sample, an explanation for the language use predictive factor in native speakers is not evident and further research is needed to explore this effect; perhaps certain domains of language use (e.g., in an academic setting) contribute more to the predictive power of the language use factor and could better elucidate this finding. Nonetheless, language use data was collected solely via self-reports, which can be problematic as there is variation in how aware and accurate language users are of their actual language use, as many speakers either over- or under-estimate their actual language use. It may be that native speakers are more accurate when self-reporting their language use, particularly, perhaps, if they are dominant in Spanish, whereas adult learners have a much harder time with accurately reporting their non-dominant language use.

6.7 Summary of findings and supporting theory

Table 38 concisely summarizes the key findings of this study and how they can be supported or explained by theory as discussed up until now.

Table 38. Summary of key findings per research question and listing of theories to support an explanation of the findings.

QUESTION	KEY FINDINGS	SUPPORTING THEORY
How do native Spanish speakers and late learners of Spanish perform with grammatical gender?	- enhanced learner performance with frequent, grammatical, and masculine noun tokens - slightly enhanced native speaker performance with high-frequency noun tokens - interactions with proficiency: advanced proficiency learners patterned *qualitatively* with lower-proficiency learner peers, yet *quantitatively* closer to native speakers - variation decreased with increasing proficiency level - higher variation for all speakers on self-paced and auditory tasks	- Complexity and Dynamic Systems Theory (Larsen-Freeman, 2012) - morphological cues (Foote, 2015; Montrul et al., 2008) - default processing strategies & biases (Eddington & Hualde, 2008; Harris, 1991)
How do the task demands impact accuracy and variation?	- intermediate and beginner learners performed better on self-paced and written tasks	- implicit vs. explicit learning mechanisms & knowledge systems (N.C. Ellis, 2005)

	- advanced learners and native speakers performed better on speeded tasks and showed no stimuli modality effects - time constraint factor most significantly predictive of performance in all speakers (although varied direction of effect)	- Declarative-Procedural Model (Ullman, 2001, 2004) - Information Processing Theory (Anderson, 1995)
To what extent can individual differences predict or account for accuracy and variation in performance?	- Spanish proficiency was the strongest predictor of learner performance, followed by metalinguistic awareness and motivation - only weekly Spanish use was a significant predictor of native speaker performance - many individual factors found to be positively correlated in learners, suggesting co-development	- Neurolinguistic Theory of Bilingualism (Paradis, 2004, 2009) - deep processing (Ehrman, 1996) - integrative motivation (Gardner, 1985) - L2 Motivational Self System (Dörnyei, 2005, 2009) - psychological theory of "possible selves" (Markus & Nurius, 1986, 1987)
Is there a multilingual effect and what is its nature?	- slight typological and quantitative multilingual learner advantage across all measures - multilingual advantage associated with presence of grammatical gender system in prior linguistic repertoire and more prominent in learners reporting 2+ additional languages - multilingualism only minimally predictive of learner performance - multilingualism NOT significantly predictive of native speaker performance	- changes to the cognitive-linguistic system: direct and indirect influence of multilingualism (Hirosh & Degani, 2018) - the role of typology & linguistic repertoire (Cenoz, 2013) - language distance (Ringbom, 2007; Jarvis & Pavlenko, 2008)

In sum, we have observed that adult learners exhibit enhanced accuracy with frequent, grammatical, and masculine noun tokens and that native Spanish speakers also demonstrate a slight increase in accuracy with high-frequency noun tokens. Proficiency modulates the relative effect of the linguistic variables on performance such that as proficiency increases, the linguistic variable effect is reduced and advanced learner accuracy scores pattern more closely with native speakers. Variation also fluctuates according to proficiency and task type such that higher proficiency is associated with less variation and self-paced and auditory tasks result in higher variation in all speaker groups. The linguistic effects uncovered in this study can be explained to a large extent by Complexity and Dynamic Systems Theory (Larsen-Freeman, 2012) regarding the role of variation in the developing linguistic system and by the role of morphological cues (Foote, 2015; Montrul et al., 2008) and default processing strategies (Eddington & Hualde, 2008; Harris, 1991). Intermediate and beginner adult learner performance was enhanced on self-paced and written tasks and these task effects also interacted with proficiency level such that advanced learners patterned very closely with their native speaker peers by demonstrating enhanced accuracy on speeded tasks and by demonstrating no stimuli modality effects. The time constraint factor was found to be most significantly predictive of both learner and native speaker performance, although the direction of this effect was reversed for advanced learners and native speakers (enhanced accuracy on speeded tasks) as compared to intermediate and beginner learners (enhanced accuracy on self-paced tasks). The task effects detected in this study can be largely explained by the theoretical contrast between implicit and explicit learning mechanisms and knowledge systems (N.C. Ellis, 2005), by the Declarative-Procedural Model (Ullman, 2001, 2004) for long-term memory systems, and by Information Processing Theory (Anderson, 1995). Spanish proficiency was found to be the strongest individual predictor of learner performance, followed by metalinguistic awareness of the target structure and motivation to learn the language. However, these individual variables were not significant factors for native speaker performance. Furthermore, many individual factors were also found to be positively correlated, suggesting that they develop in unison throughout the language acquisition process. The impact of individual factors on performance can be largely explained through an examination of the Neurolinguistic Theory of Bilingualism (Paradis, 2004, 2009), integrative motivation (Gardner, 1985), the L2 Motivational Self System (Dörnyei, 2005, 2009), as well as by the psychological theory of "possible selves" (Markus & Nurius, 1986, 1987). Finally, a slight typological and quantitative multilingual learner advantage was detected across all measures and was found to be associated with the presence of a grammatical gender system in the prior linguistic repertoire of the adult learner and this observed multilingual advantage was also found to be more prominent in learners reporting two or more additional languages. However, according to linear regression modeling, the multilingualism factor was found to be only minimally predictive of late learner performance and was not significantly predictive of performance in native speakers. The multilingual learner advantage could be explained in part by changes to the cognitive-linguistic system through both direct and indirect influence of multilingualism (Hirosh & Degani, 2018), the role of typology and linguistic repertoire (Cenoz, 2013), and by language distance (Ringbom, 2007; Jarvis & Pavlenko, 2008). We will now move our discussion to the question of *learnability*.

6.8 Is grammatical gender acquirable in adult learners?

Now that we have considered collective predictive modeling to identify the independent factors that can predict performance with grammatical gender in both native speakers and nonnative learners at different proficiency levels, an essential question remains on the issue of *learnability* in late/adult language acquisition: What can the present findings tell us about whether or not grammatical gender can be (fully) acquired by late/adult learners and, furthermore, how can the factors that we have uncovered here explain this outcome or elucidate the question of *learnability*?

Both generativist and cognitivist accounts of language acquisition acknowledge that implicit knowledge—and <u>not</u> knowledge *about* language—crucially comprises *linguistic competence* such that language acquisition is evident in what language users know *intuitively* (White, 1989, 2003; Gregg, 1989; Krashen, 1985). Therefore, is grammatical gender acquirable *exclusively* through an implicit process in late acquisition? The present findings suggest that while gender appears to be (mostly) acquired in advanced proficiency adults, the trajectory of acquisition also appears to diverge from exclusively intuitive and implicit acquisition to the extent that adult acquisition is scaffolded by explicit knowledge, such as metalinguistic awareness of the target structure. Explicit knowledge scaffolding to compensate for deficits in implicit linguistic competence in the developing mental grammar of the adult language learner is particularly evident at beginner levels of proficiency, but the present findings also seem to indicate that as proficiency increases, the need for explicit knowledge scaffolding diminishes; ultimately, advanced instructed adult learners pattern very closely to their native speaker peers—demonstrating enhanced performance on speeded tasks that tap into implicit linguistic competence—while still maintaining high levels of metalinguistic/explicit knowledge observable during a metalinguistic exit survey inquiring about the nature of the study and target structure. The co-development of implicit linguistic competence scaffolded by explicit linguistic knowledge about the target language is likely to be more prevalent in adult learners in the instructed context, as is the case with the present learner sample.

Although the feature *gender* could be represented at the abstract level in the mental grammars of late learners, its retrieval or the manner in which this feature is accessed for language processing, could vary in different language users such as in early as compared to late learners. Recall that Prévost and White (2000) argue that late learners do have the abstract gender feature represented in their L2 grammars, but that errors still occur nonetheless due to an assembly or production problem, essentially, a *computational difficulty* during online language processing, as they posit in the Missing Surface Inflection Hypothesis. Empirical research demonstrating higher accuracy rates in late learners during cognitively offline production tasks than during online and unplanned tasks (e.g., Gamboa, 2012; Grüter et al., 2012; Alarcón, 2011; Montrul et al., 2008) provides support for their hypothesis. In research comparing heritage learners exposed to Spanish since birth and traditional L2 adult learners exposed to Spanish post-puberty, Montrul et al. (2008, 2016) argues that both learner groups know something about grammatical gender in Spanish, but that this knowledge or competence might be stored, represented, and reproduced differently in the mental grammars of early-exposure and late-exposure learners. Alarcón (2011) further instantiates this distinction by positing that post-critical period (adult) learners are more susceptible to computational deficits observed in spontaneous production of gender agreement than those who acquired the target language since birth, such as early-exposure heritage learners and native speakers. Nonetheless, it is unclear how a 'computational difficulty' or 'deficit' during online language processing may relate to the declarative-procedural long-term memory distinction; it seems that what could be interpreted as a 'computational difficulty' could have its source in the distinct memory systems involved such that difficulties during online language processing, conditioning the use of implicit linguistic competence, would necessarily reflect deficits in the representation of implicit language knowledge in procedural memory and indicate that, if the gender feature is still represented mentally, it must be represented to a larger extent in the form of explicit knowledge about language stored in declarative memory, better suited to be accessed during offline processing. However, this distinction is likely more complex than what it appears, and it might be that the different subcomponents of gender are represented differently in both procedural and declarative memory. For example, perhaps the syntactic agreement component is stored in procedural memory as implicit, whereas the gender assignment component and corresponding gendered articles that accompany the noun may be governed to a larger extent by explicit knowledge of the lexicon stored in the declarative long-term memory system. These differences, then, would become observable in linguistic behavior through varying accuracy rates not only from different task constraints such as speeded as compared to self-paced performance tasks, but also along the lines of the linguistic features inherent to grammatical gender, such as the degree of transparency of the noun morphology and the relative frequency of the target noun and its corresponding derivational morphemes that may modulate knowledge of gender assignment, during tasks that variably condition online and offline language processing.

Moving beyond the features of grammatical gender alone, we can frame the argument of *learnability* around the concept of ultimate attainment (UA) in late acquisition research. Although the majority of UA studies use highly educated participants as a native control group to which to compare adult learners (Andringa, 2014), a very different picture emerges from studies that use a native control group that includes lower socioeconomic status speakers and these studies tend to show more adult learners performing within the native speaker range (e.g., Andringa, 2014; Dąbrowska & Street, 2006; Hulstijn, 2015). In this sense, there are two principled ways of characterizing native speakers: 1. *shared basic language cognition* (see: BLC framework, Hulstijn, 2015); and 2. *extended* or *higher language cognition*. While BLC pertains to the frequent units and constructions that are shared by all native speakers, extended language cognition pertains to the infrequent units and constructions that distinguish native speakers from one another along extralinguistic dimensions such as socioeconomic status, level of education, and print exposure (Hulstijn, 2015). In fact, print exposure and level of education are predictors of attainment that are largely similar for native and nonnative speakers alike, although not necessarily equally weighted for both speaker groups (Dąbrowska, 2019). Therefore, from the findings of the present study, we argue here that frequent and canonical forms of grammatical gender are, in fact, acquirable for late learners and that the linguistic factors that condition performance in late learners also condition performance, albeit to a lesser extent, in adult native speakers too. Furthermore, at advanced levels of nonnative language proficiency, adult learners pattern very closely with their native speaker counterparts regarding the relative effect that the linguistic variables have on performance. Nonetheless, more diverse sampling of native speakers is needed by which to compare late learners to native speakers of different socioeconomic and educational backgrounds. From the present sampling of educated native speakers and adult learners, it appears that gender is (fully) acquirable at advanced levels of proficiency, although individual and group variation exists to a larger extent in nonnative speakers, yet we may see a different outcome if speakers of lower socioeconomic and educational backgrounds were compared and if nouns exhibiting exceptional morphology and highly infrequent nouns were included in the task stimuli.

Previous empirical findings show that even advanced proficiency late learners display persistent errors with gender agreement, primarily in more spontaneous oral production (Montrul et al., 2008). However, research comparing heritage learners exposed since birth to Spanish and late/adult learners of Spanish exposed only after puberty suggests that Spanish gender agreement is, in fact, acquirable *irrespective of the age of acquisition* as both early and late exposure learner groups were found to demonstrate systematic errors in their performance (Alarcón, 2011; Montrul et al., 2008), in accordance with their respective proficiency level. Therefore, the essential question in generativist terms becomes whether highly proficient late learners can ever acquire the parametrized functional feature of grammatical gender if it is not already instantiated in the learner's native mental grammar (Spino-Seijas, 2017). Or in other words, can adult language learners integrate into their implicit linguistic competence a grammatical structure that is not already instantiated in their native language grammar? Our findings, especially with regards to task effects, show that advanced proficiency late learners have very comparable performance to their native speaker counterparts by performing better on speeded tasks as compared to self-paced tasks and by not exhibiting input modality effects. Moreover, advanced and intermediate late learners actually outperformed their native speaker peers on self-paced tasks. Therefore, we can conclude from these findings that late learners do, in fact, have the grammatical gender feature instantiated in their mental grammars at an abstract and implicit level as demonstrated during online/speeded language processing, and furthermore, also have grammatical gender represented at a more explicit level in terms of declarative and metalinguistic knowledge that can be operationalized to their advantage for offline/self-paced language processing. Intermediate and beginner learners, on the other hand, performed significantly better on self-paced tasks, suggesting that their performance with grammatical gender is most supported by explicit/declarative knowledge available during offline processing whereas their advanced learner and native speaker peers appear to depend more on implicit linguistic competence observable to a larger extent during online language processing under a time constraint. In sum, the gender feature is, in fact, acquirable by late learners, but its representation and relative distribution in the form of explicit knowledge and implicit competence varies in function of overall proficiency in the language.

When considering the question of *learnability* in late/adult language acquisition, it is essential to examine the highest proficiency, near-native adult learners whose grammar may be outwardly indistinguishable from a native speaker grammar. Admittedly, these types of learners, while existent, are hard to find and the present thesis research certainly suffered from a lack of the highest proficiency adult learners as only fourteen of the recruited participants tested at the advanced level of proficiency. Nonetheless, the present and limited findings seem to provide preliminary evidence that advanced proficiency late learner grammar is largely indistinguishable from that of native speakers. This finding for advanced proficiency late learner performance largely corroborates the work of Bruhn de Garavito (1999) who found that near-native adult learners of Spanish (from both L1 English and L1 French backgrounds) were, in fact, able to acquire subtle grammatical properties that were not evident in the input, therefore pointing to implicit modes of language acquisition in adult learners. Our findings also corroborate those of White et al. (2004) who found that intermediate and advanced learner performance with grammatical gender in Spanish was not significantly different from the native Spanish speaker control group, and the effect of L1 and prior linguistic experience were non-significant in their sample of L1 English and L1 French late learners of Spanish. Nonetheless, our findings diverged slightly from those of White et al. (2004) since only advanced learners—and not intermediate level—patterned with native speaker performance in our sample. Perhaps the relative difficulty of the tasks used to assess performance could explain this. Nonetheless, from a generativist framework, the fact that high-proficiency adult learner performance is not significantly different from native speaker performance suggests that late language acquisition can also be constrained to some extent by the implicit language-specific and innate learning mechanisms of UG.

Relating the learnability question back to different theoretical accounts, recall that some researchers argue that there is a so-called *representational deficit* or *maturational constraint*—the so-called "learnability problem"—for adult language acquisition (e.g., Morphological Underspecification Hypothesis, McCarthy, 2008; the Fundamental Difference Hypothesis, Bley-

Vroman, 1989, 1990) while others support the view of *full-access* (e.g., Missing Surface Inflectional Hypothesis, Prévost & White, 2000) to the implicit and innate learning mechanisms of Universal Grammar (UG), in which errors are merely the result of a “mapping problem” in nonnative language production or a “computational difficulty” during online language processing (Spino-Seijas, 2017; López Prego, 2015; Alarcón, 2011; Montrul et al., 2008). Representational deficit accounts posit that adult language learners are forced to rely on explicit or general learning mechanisms to acquire another language post-puberty and do not maintain access to the implicit and linguistically specific learning mechanisms of UG. More specifically, the Morphological Underspecification Hypothesis posits that gender errors may be more common with one gender than the other due to the overgeneralization of a default form. Often times, the masculine gender class is treated as default, as previously discussed; however, this does not necessarily constitute evidence of a fundamental representational deficit. Although our empirical findings seem to suggest that adult learners, particularly at the beginner and intermediate levels of proficiency, depend to a larger extent on their explicit and metalinguistic knowledge of grammatical gender observable in their relatively enhanced linguistic performance during cognitively offline self-paced tasks and on the metalinguistic awareness exit survey, thereby pointing to a deficit in their implicit linguistic competence and therefore the need to depend to a greater extent on explicit knowledge, this effect subsides at more advanced levels of proficiency in which advanced learners pattern with native speakers in this regard. Furthermore, although late learners are much more susceptible to demonstrating enhanced accuracy with unmarked forms (masculine, transparent), native speaker performance was also found to be affected by the linguistic variables analyzed, particularly relative noun frequency, indicating that processing strategies are present in both language user groups. If these processing strategies are instantiated in both native and nonnative learners alike, it stands to reason that modulating accuracy levels for certain forms cannot alone be concrete evidence of a representational deficit; those learners with a representational deficit at lower levels of proficiency would necessarily use general cognitive processing strategies to a larger extent to compensate for this deficit, yet the use of these strategies alone cannot by itself implicate a representational deficit.

More compelling evidence for a fundamental representational deficit would be in varied performance according to task demands that condition online and offline language processing, involving compensation or scaffolding via explicit knowledge. Nonetheless, our findings suggest that the task demands effect subsides at advanced levels of proficiency, indicating that the relative balance between explicit language knowledge and implicit linguistic competence fluctuates along the natural trajectory of late language development, potentially arriving at native-like competence, yet through much more extensive explicit knowledge scaffolding than during early L1 development. Does this necessarily indicate a fundamentally different process of language acquisition? Perhaps, and particularly with adult language learners in the instructed context. But does this also mean there must be a fundamental representational deficit in final attainment or end state for adult learners? No, not entirely. Although we cannot be sure from external linguistic behavior alone, it seems that adult language users are fully capable of acquiring the gender feature if they arrive at advanced and native-like levels of proficiency, yet their path or developmental trajectory to a native-like end state seems to be distinct from that of a native speaker who acquired the language since birth and therefore was not able to nor needed to use explicit and metalinguistic knowledge to the same extent. Therefore, the learnability problem is more of an issue of *trajectory* or, rather, a question of *how* gender is acquired rather than *whether* gender can be acquired in late acquisition.

6.9 Limitations

This study has several important limitations that should be highlighted in order to more accurately interpret the findings as well as in order to continue to improve upon and refine the research methodology and analysis for future research into factors that affect native speaker and adult learner performance.

One important limitation of this study concerns the task stimuli. Some of the target prompts that participants had to respond to were inadvertently unnatural, and therefore, participants' correct/incorrect responses could have been more related to the unnaturalness of the utterance than strictly in response to the grammaticality of the prompt, which was the intention. For example, in Task 2 (speeded oral imitation task), one of the ungrammatical target prompts is "*cada queso es negra*" (each cheese-*masc.* is black-*fem.*). Even some native speakers were confused by how unnatural this phrase sounded. Furthermore, two different copular verbs were utilized in many of the target prompts: /*estar*/ ('*to be'* for temporary conditions) and /*ser*/ ('*to be'* for permanent states and general characteristics). At lower proficiency levels, learners might have been confused by the focus of the task (which was intentionally not made explicit to them) and consequently their grammaticality judgement might have been more related to their assessment of the correct use of the copular verb /*ser*/ or /*estar*/ than it was in reaction to the grammatical gender syntactic agreement component of the target sentence, leading some learners to evaluate a given prompt as ungrammatical, even though the phrase was grammatically correct for gender agreement. This suspicion was later confirmed for several beginner participants who when asked during the metalinguistic awareness exit survey what they thought the aim of the study was included "/*ser*/ vs. /*estar*/" in their response. It also doesn't help that this verb distinction is explicitly taught and regularly reviewed in beginner Spanish courses and so beginner learners are sometimes hyper-vigilant with regards to the use of the /*ser*/ vs. /*estar*/ copular verb distinction. Therefore, our interpretation of beginner learner accuracy on these prompts could potentially be confounded by this confusion. Nonetheless, only 15 target prompts contained the /*estar*/ copular verb instead of the copula /*ser*/ out of 104 prompts total and 16 learner participants (approx. 14% of all participants) mentioned the "/*ser*/ vs. /*estar*/" distinction during their exit survey. Therefore, in future research, it is advisable to maintain constant the copular verb (either /*ser*/ or /*estar*/, but not both), particularly when working with beginner-level participants, so as to avoid this potential confusion.

Regarding the analysis of the effect of noun morphology in the present study, only the most typical/canonical overt vs. non-overt noun morphology distinction was examined between nouns that follow the prototypical pattern of feminine nouns ending in /*-a*/ and masculine nouns ending in /*-o*/ as compared to Spanish nouns that do not follow this pattern. Although the descriptive results of the present study do corroborate the idea that learners use this prototypical Spanish noun morphology pattern when processing noun gender assignment in Spanish, it is likely that learners also use other noun morphology patterns beyond this canonical distinction when processing gendered noun morphology. For example, adult learners may be sensitive to other Spanish noun morphology patterns such as feminine nouns that end in /*-dad*/ (e.g., *universidad*-fem, "university") or /*-ción*/ (e.g., *habitación*-fem, "bedroom") as these suffixes tend to be quite systematically feminine in Spanish. The present dataset could be reexamined to compare learner accuracy rates on different Spanish noun suffixes and further examine this pluralistic view of the noun morphology effect at different proficiency levels as learners' sensitivity to noun morphology may be tuned to certain patterns to varying degrees of influence at different proficiency levels. For example, perhaps beginner learners are more sensitive to the prototypical binary < /-*a*/ vs. /-*o*/> pattern, but intermediate proficiency learners start to diversify morphological cues when processing gender assignment, becoming sensitive to other gender-revealing patterns in Spanish noun suffixes.

The uneven and relatively small (in some cases) groupings in each individual variable group constitute another important limitation of this study, particularly for the multilingualism analysis in examining the effect of prior linguistic repertoire and the effect of the number of additional languages reported. A more diverse sample of multilingual learners is needed who report one, two, and three additional languages that are also varied typologically for both the presence and type of grammatical gender, that is, recruit more multilingual learners who do not have binary grammatical gender present in their prior linguistic repertoire. Another limitation regarding the effect of multilingualism is the impact of proficiency in all languages reported. Although Spanish proficiency was tested, proficiency in additional reported languages was not tested nor controlled for. For a more in-depth and valid analysis of the effect of multilingualism and prior linguistic repertoire, it would be essential to have some sort of objective measure of proficiency beyond self-report in multilingual learners' additional languages reported as this would likely influence the relative effect that knowing additional languages has on performance with L*n* Spanish. In other words, if a learner only has beginner knowledge of an additional language that also exhibits binary grammatical gender, it is likely that this would afford much less (if any) advantage for performance with grammatical gender in Spanish than someone who is fully proficient in another binary gendered language.

More specifically, although a concerted effort was made to recruit instructed learners from different proficiency levels, tested Spanish proficiency scores revealed that the vast majority of the learner sample was skewed towards the beginner level of proficiency, producing unbalanced proficiency groupings: sixty-two beginners and only fourteen intermediate and fourteen advanced learners. It may be that within the context of instructed language learning, it is necessary to recruit more learners from graduate-level courses as well as learners who have spent time abroad in the target language community, thereby increasing the likelihood of recruiting participants at the most advanced levels of proficiency. Nonetheless, this would imply a more diverse sample that includes naturalistic and instructed exposure to the target language.

Furthermore, the present study design did not record reaction times (RTs) and therefore there is no objective way to confirm to what extent participants may have varied in how quickly they reacted to speeded tasks; in other words, the task conditions were intended to be uniform for all participants, but in reality, participants likely varied in how quickly they responded and some participants may have taken a bit more time to think about prompts than others. Although task completion times were recorded based on the time of each audio recording in *Audacity*, completion times were ultimately excluded from the analysis as multiple delays and recording anomalies were detected in which certain participants paused the task to ask questions or make a comment, whereas others moved through each task as quickly as possible with no interruption. Furthermore, due to the online nature of the study session via Zoom, there were some connectivity issues observed in which participants' connection speed would become slower and therefore there was a delay in what the participant saw or heard, thereby increasing their task completion times in an unproductive and uninterpretable manner.

A final limitation of this study pertains to the socioeconomic and academic background of both the native Spanish speakers and adult learners recruited. All participants were, or had recently been, university students, either at the undergraduate or graduate level. It is likely that both learners and native speakers alike who possess diverse academic backgrounds and literacy levels would be differentially affected by the linguistic variables, task demands, and individual factors considered in the present study. For example, metalinguistic and metacognitive awareness are likely more prevalent among formally educated and highly literature adults. Furthermore, among the recruited learners, all participants were instructed learners. Although this was a major design feature of the study—to look exclusively at language acquisition in the instructed context—more research is emerging that indicates significant difference between informal/naturalistic and formal/instructed language learners (See: R. Ellis, 1989; C. Muñoz, 2008; Alptekin, 2007). Therefore, it is likely that the linguistic variables, task conditions, and individual factors examined would have varying effects if different learning contexts were compared. Moreover, comparing naturalistic and instructed learners would be more ecologically valid in the sense that today many multilingual language learners are *casual* learners, picking up some language skills here and there through their friends and social interactions, through travel, for professional reasons, or simply for pleasure through media consumption. This seems to be more the case in Europe, for example, than in North America, but nonetheless the world has become more globalized since the traditional (monolingual anglophone) L2 instructed learner was originally established as the norm in language acquisition research. Nonnative language acquisition research only stands to benefit by taking a broader perspective on exactly who constitutes a qualified study participant—considering language, socioeconomic, and academic backgrounds as well as learning context.

6.10 Future Work

Based on the findings and limitations of this dissertation research, avenues of future research will be discussed in order to further the aim of identifying and exploring the most important linguistic, task, and individual factors that can predict or explain the linguistic performance of native speakers and nonnative language learners.

The nature of the task effects detected in this study could be examined more closely and precisely by teasing apart the effect of stimuli presentation modality from the effect of a time constraint; the present study did not include a task that presented aural stimuli in a self-paced manner (i.e., allowing participants to listen to the audio as many times as needed before advancing to the next prompt). Therefore, the present study cannot draw conclusions about the effect of aural stimuli that is independent from the effect of a time constraint. In future research, a self-paced aural stimuli task could be administered and accuracy scores on this task could be compared to accuracy on other task types that control for one task condition variable at a time: 1.) a speeded auditory task; 2.) a self-paced auditory task; 3.) a speeded written task; and 4.) a self-paced written task. In this manner, scores on written tasks could be compared by task timing (i.e., speeded vs. self-paced) and scores on auditory tasks could also be separately compared by task timing in order to determine the extent to which these task variables may produce independent effects on performance.

In addition, the direction of influence between different learner factors may be investigated through longitudinal studies of learners by surveying their motivation levels, attitudes, metacognitive and metalinguistic awareness, and language use throughout the language learning process (i.e., at different levels of proficiency) as they progress. Furthermore, the different subcomponents of individual factors could be examined, including motivation (motivated learning behavior, ideal L2 self, ought-to L2 self, and motivated learning experience), attitudes (feeling like oneself when speaking the target language, identifying with a Spanish-speaking culture, desire to use Spanish as a native speaker, and a desire to be perceived by others as a native speaker), metacognitive awareness (knowledge about and regulation of cognition), and language use in different contexts (with friends, with family, at school/work, when engaging in self-talk, and when counting) to explore the relationship between these subfactors and accuracy and proficiency scores as this type of inquiry would help to further elucidate the construct of each individual factor and determine if certain subcomponents may be more important or predictive of performance than others. Further analysis is also needed regarding how proficiency may interact with individual learner variables. In other words, could the effect of individual learner variables like metacognitive awareness and attitudes simply be explained by correlations with proficiency? Future research could examine the effects of each learner variable per proficiency level to examine if beginner, intermediate, and advanced learner performance is equally affected by these individual factors. Perhaps these individual factors do not have as much (or any) predictive power at the advanced proficiency level but have a larger effect size at the beginner and/or intermediate levels of Spanish proficiency.

Another important point that the present study did not address is how the effect of the linguistic variables (noun gender class, noun morphology, domain of agreement) interact with the individual factors (motivation, metacognitive awareness, metalinguistic awareness, attitudes, target language use, prior linguistic repertoire) to determine how learners with different individual characteristics may be differentially affected by the linguistic features of the target structure. For example, are more metalinguistically and metacognitively aware learners more sensitive than their less aware peers to noun morphology? This line of research has the potential to elucidate the nature of how the linguistic structure in question is interpreted, processed, and produced by learners with varying characteristics.

Future research could also dive more deeply into the multilingual analysis by collecting more data on multilingual learners from different language backgrounds examining the effect of language typology (grammatical gender type, asymmetry analysis of gender assignment across languages) and language learning context (informal/naturalistic vs. formal/instructed) of the additional reported languages of multilingual learners. In addition, future research into the exact nature of the multilingual effect observed in the present study could examine the age of acquisition (AoA) of the additional languages reported by multilingual participants, particularly the AoA of other gendered languages, as early acquired additional gendered languages may have more of an impact on the acquisition of L*n* grammar than more late acquired additional languages. Finally, proficiency level should be tested in all the additional languages reported by multilingual participants as this may modulate the effect that a multilingual background has on the learning of a novel language (i.e., known languages at lower proficiency levels are perhaps less likely to produce a multilingual learner advantage).

Furthermore, frequency effects could be further investigated as both token and type frequency (see: N.C. Ellis, 2002) and per proficiency level as frequency effects are likely to be modulated to varying extents by learners' overall proficiency level—learners become more or less sensitive to frequency as their proficiency develops and the direction of frequency effects may change too (i.e., beginner learners could be more sensitive to high-frequency items whereas advanced learners may be more sensitive to infrequent items). In addition, frequency by regularity interactions could be examined, as uncovered by N.C. Ellis and Schmidt (1997), in which canonical and noncanonical noun targets that are classified as frequent and less frequent may be compared so as to determine to what extent the present sample may replicate the findings of previous frequency research with regards to the dynamic interaction between token/type frequency and regularity/systematicity of the item (in this case, overtly or non-overtly marked gendered nouns).

Moreover, target language use could be examined more closely by considering separately each dimension/context of target language use (at school/work, with friends, with family, during self-talk, when counting) as an independent variable to explain to some extent variation in performance among nonnative learners of Spanish; De Carli et al. (2015) assert that further research is needed to explore language use in greater detail as they found a significant effect of language use classified bimodally, but argue that the detected language use effect may be more complex and, therefore, not fully represented with this bimodal categorization. Future research could also examine how language use in different contexts may interact with tested and self-reported Spanish proficiency as more proficient learners are likely to use the target language more in certain contexts and, likewise, using the target language more in certain contexts may lead to proficiency gains; the direction of influence of the language use factor could be explored through longitudinal studies that follow language learners' self-reported language use as their tested proficiency and instructional time increase (e.g., over the course of an academic year).

Finally, future research should include reaction time (RT) data as a dependent response variable, in addition to accuracy and variation, in order to draw more precise conclusions regarding the processing mode of language learners and native speakers during varied task demands. In other words, the task conditions were meant to be uniform for all participants; however, certain participants took more time than others to think about the prompts and or stalled in their response by asking questions or making a comment. This varied type of behavior could have important consequences when drawing conclusions regarding what knowledge type and memory systems are being deployed by language users during different tasks; participants who took more time to respond, despite the speeded task instructions to respond as quickly as possible, might have been able to access more of their explicit linguistic knowledge stored in declarative memory than their peers who responded more quickly, thereby indicating the ability to fluently access implicit linguistic competence stored in procedural memory. The tasks were designed to variably condition or favor the use of either explicit (self-paced, written) or implicit (speeded, auditory) knowledge; however, participants necessarily responded differently to these task conditions and might have, to some extent, strategically modulated their behavior to compensate for deficiencies (by responding more slowly to have more time to think and access explicit representations of language) or exploit their (implicit) linguistic competence (by responding more quickly). In particular, it would be informative to compare RT data to accuracy score data on each task type and per proficiency group in order to examine how language users may variably access implicit and explicit knowledge, how this may change with increasing proficiency, and how the task constraints affect reaction times per proficiency group. It could be, for example, that RT and accuracy are positively related at the beginner proficiency level, yet are more inversely related at more advanced proficiency levels, which is what the findings of the present study seem to indicate in native speakers and advanced learners; their accuracy rates actually went down when they had more time to think about the prompt whereas when they were instructed to respond as quickly as possible, their accuracy rate increased, likely because they were accessing to a larger degree their implicit linguistic competence, which is less variable and more systematic. In sum, it is difficult to draw extensive conclusions regarding explicit-implicit language knowledge and declarative-procedural memory systems when reaction time data is not considered. Speed of response may be the best indicator of which knowledge type and memory system are being accessed at any given time in the mental grammar of the language user. Nonetheless, we can still draw some tentative conclusions from the present dataset regarding how task conditions are *likely* to condition the use of certain knowledge types and memory systems, although we cannot be sure.

In the following chapter, the aims and key findings of this dissertation research will be briefly revisited. Furthermore, the significance and contributions of this study to the larger fields of adult language acquisition and multilingualism will be discussed and some possible pedagogical implications of these findings will be proposed.

Chapter 7: Conclusion

In this thesis, some key factors have been explored that influence both native speaker and adult learner performance with a grammatical structure identified as problematic for the late acquisition of morphosyntax. This exploration has been oriented by different types of factors that are believed to variably influence language performance: speaker/learner-*external* factors including linguistic features inherent to grammatical gender (gender class, noun morphology, domain of agreement, noun frequency), task demands in terms of the stimuli modality (auditory vs. written) and the time constraint factor (speeded vs. self-paced), and speaker/learner-*internal* factors that may influence how one acquires, processes, accesses, and uses language, including target language proficiency, motivations to learn the language, attitudes about the target language and target language community, average weekly language use, metacognitive awareness of the individual's learning process, metalinguistic awareness as explicit knowledge of the target structure, and the influence of a multilingual background as well as the typology of one's prior linguistic repertoire. The extent to which these explanatory factors may predict linguistic performance with grammatical gender in Spanish have been examined in terms of both accuracy rates and variation at the individual (intra-learner/speaker variance) and group (inter-learner/speaker standard deviation) levels. The key objective was two-fold: 1. to understand what factors may predict performance in different language user groups (native speakers and late learners), and 2. to draw principled conclusions about what knowledge types—including implicit and explicit—and memory systems—including declarative and procedural long-term memory—language users draw upon when processing and producing language and how this may be different for learners of varying proficiency levels and of varying characteristics. Through an examination of knowledge and memory system types and their relative proportions in different language users, we can better address the underlying question of *learnability* in late acquisition: to what extent can adults acquire novel language grammar and how may their developmental trajectory in the target language vary from their native speaker peers? This is the essential question I hope to have addressed here.

We will now take a moment to briefly revisit the key findings of this thesis research as they pertain to the linguistic, task, and individual factor effects uncovered. With regards to the inherent features of grammatical gender, the present findings indicate enhanced late learner performance with frequent, grammatical, and masculine noun tokens and a slightly enhanced native speaker performance with high-frequency noun tokens as well. Interactions with proficiency were also detected such that advanced proficiency learners patterned *qualitatively* with their lower-proficiency learner peers in terms of the relative effect of each linguistic variable yet were *quantitatively* more similar to their native speaker counterparts in terms of accuracy rates. Variation in performance was also found to decrease with increasing proficiency level and relatively higher variation was detected in all participants on self-paced and auditory task types. With regards to the impact of varied task demands, intermediate and beginner learner performance was found to be enhanced on self-paced and written tasks whereas advanced learners and native speakers performed better on speeded tasks and showed no stimuli modality effects. Regarding the impact of individual factors, tested Spanish proficiency produced the most differentiating effect between late learners followed by their metalinguistic awareness level, and finally their motivation to learn the language proved to be important in the initial descriptive findings and individual correlations. Individual factors were largely not found to be important for native speaker performance as only a slight effect of weekly use of Spanish was detected for native speakers, however the exact nature of the language use effect in native speakers remains to be explored. Finally, a slight typological and quantitative multilingual learner advantage was detected across all measures and this advantage was found to be associated with the presence of a grammatical gender system (of any subtype) instantiated in the prior linguistic repertoire of the multilingual learner. Furthermore, the multilingual advantage appeared to be more prominent in learners reporting knowledge of two or more additional languages beyond English and Spanish. Again, similar to individual factors in general, a multilingual background appears to have an impact on learner performance, however this is not the case for native speakers.

Key findings from comprehensive linear regression modeling indicate that token grammaticality, noun frequency, and task processing type (i.e., speeded online processing vs. self-paced offline processing) are all significant and independent learner-external predictors of performance accuracy when all late learners are considered together and in beginner learners separately. Furthermore, metalinguistic awareness of grammatical gender was found to be a significant learner-internal predictor in all learners and in beginners. However, none of these factors were independently predictive of performance in intermediate and advanced proficiency learners. This is a finding that should not be overinterpreted given that the vast majority of the learner sample tested at the beginner proficiency level and, therefore, with a larger sample size of intermediate and advanced proficiency learners, it is possible that these predictors may reach significance, as the initial descriptive findings had suggested, particularly in intermediate learners. However, the relative effect sizes compared per learner proficiency group seem to indicate that these individual factors become much less predictive of performance as proficiency increases. Nonetheless, significant predictors of performance were also found for the native speaker group. The speaker-external factors of noun frequency and task processing type were also found to be significant predictors of native speaker performance with grammatical gender in addition to the speaker-internal factor of average weekly Spanish use. No other factors were found to be significant predictors of performance in native speakers according to comprehensive linear regression modeling.

This study makes several novel contributions to the fields of late language acquisition research and multilingualism. First, evidence has been provided of how adult learners and native speakers alike are variably affected by both speaker-external and speaker-internal factors and show how both these language user groups demonstrate variation in their performance that is conditioned by task type, proficiency level, and frequency of the stimuli. The present findings challenge the prototypical image of adult native speakers as an invariable and homogenous group; this thesis research demonstrates that native speakers too vary in their performance and that this variation can be modulated by the demands of the task and the relative frequency of the stimuli. This study further elucidates how late learners may differentially draw upon their explicit language knowledge stored in declarative memory and their implicit linguistic competence stored in procedural memory, as evidenced in participants' performance with varied task demands designed to differentially condition different linguistic knowledge sources. Crucially, the present findings and analysis provide evidence of how the relative balance between these knowledge sources exploited during linguistic performance is modulated by the late learner's global proficiency level in the target language and also by the particular demands of the task at hand.

This thesis research also contributes to our understanding of the individual differences between late learners that can impact performance and explain to a large extent the immense variation observed in adult language learner performance; although the initial descriptive findings uncovered a variety of individual learner factor effects, through inferential analysis using multiple linear regression modeling, the present sample demonstrates how differences in learner performance can be largely attributed to global tested proficiency level, differences in metalinguistic awareness, and motivational orientation. This thesis further extends the domain of inquiry on the impact of individual differences through an analysis of the interaction between task effects and individual factors, providing important data on how motivation, attitudes, metalinguistic awareness, metacognitive awareness, and language use variably impact performance according to the linguistic knowledge type (explicit/implicit) and memory system (declarative/procedural) conditioned. For example, evidence has been provided that more metalinguistically aware individuals—including both native speakers and adult learners alike—tend to perform better on tasks in which they can take their time, likely tapping into their explicit knowledge stores in declarative long-term memory, and the present analysis has shown how learners who exhibit a more positive motivational orientation tend to perform better than their peers on tasks posited to tap more into their implicit linguistic competence in procedural memory (i.e., speeded and auditory tasks). Therefore, from these individual factors per task type effects, we can observe how individual differences can variably impact access to different sources of linguistic knowledge, namely, that motivation has more of an effect on implicit linguistic competence whereas metalinguistic awareness has more of an impact on explicit language knowledge. Moreover, this thesis makes a novel contribution by demonstrating how different individual factors are correlated with one another and therefore may develop in unison throughout the language acquisition process; for example, Spanish proficiency was shown to be highly and significantly correlated with Spanish attitudes and motivation, indicating that as the adult learner develops more competence in the target language, their positive attitudes about the target language community increase as does their motivation to learn the language.

Furthermore, this thesis research makes a novel contribution to the field of multilingualism, and, more specifically, to the body of research on how a multilingual language background can affect subsequent language learning in adults by examining performance with a particular grammatical structure. Evidence was found for a typological and quantitative multilingual learner advantage that is instantiated across all task measures but is most pronounced on tasks that condition implicit linguistic competence (i.e., speeded and auditory tasks) and on the more difficult token type of ungrammatical feminine nouns, a stronger indicator of acquisition of grammatical gender. Crucially, the findings of this thesis research provide evidence that the multilingual learner advantage appears to be associated with the typology of the learner's prior linguistic repertoire such that knowing another gendered language affords greater advantage than knowing a non-gendered additional language. In addition, the effect of multilingualism appears to be quantitative in nature such that knowing more than two additional languages seems to afford more advantage than knowing just one additional language. In sum, this thesis contributes to our understanding of how a multilingual background influences novel language acquisition by elucidating the nature of the multilingual advantage as typological, quantitative, and associated with enhanced implicit linguistic competence. In contrast, although initial descriptive findings provided some evidence for a multilingual advantage in native speakers on self-paced tasks only, subsequent inferential analysis suggests that multilingualism does *not* have a significant impact on performance with Spanish as a native language, which to my knowledge, had not been previously examined in the multilingualism literature.

Finally, this thesis contributes to the growing body of late language acquisition research by providing an analysis of how the linguistic features of a particular structure impact grammatical processing and performance by providing evidence that late learners show significantly more accurate performance with grammatical and high-frequency tokens. Likely the most notable contribution of this thesis to the larger field of adult acquisition research pertains to the present analysis of how the relative impact of linguistic, task, and individual factors changes as the late learner develops proficiency in the target language. That is to say, these factors are not static in nature, but rather modulate performance in accordance with the developing linguistic system, akin to compensation strategies or scaffolding of learning. This thesis provides evidence that adult learners use explicit language knowledge through metalinguistic awareness of the target structure to scaffold their performance, observable in enhanced performance on the metalinguistic awareness exit survey, relative to general performance, and observable in enhanced performance on self-paced tasks in more metalinguistically-aware language users; in other words, both late learners and native speakers—to a lesser extent—strategically exploit their explicit language knowledge when they are provided sufficient time to do so. In addition, the present findings contribute to the body of research on native and nonnative language processing by providing evidence of how sensitivities to linguistic features and task constraints diminish with increasing global proficiency such that clear qualitative and quantitative differences emerge in learners' relative sensitivities to these factors as their proficiency develops. The present research has shown that advanced-proficiency late learners pattern qualitatively with their lower-proficiency peers in terms of sensitivity to the inherent linguistic features of grammatical gender, exhibiting more accurate performance with grammatical, masculine, high-frequency, overt noun tokens, yet quantitatively their accuracy scores are closer to their native speaker counterparts than to their intermediate-proficiency learner peers. Furthermore, this study provides evidence that advanced learners' sensitivity to task effects mirrors native speaker sensitivities such that advanced learner performance is also enhanced on time-constrained tasks and no modulation in accuracy is exhibited for stimuli modality differences. In sum, this thesis contributed to the domain of native and nonnative language processing by providing evidence that as global proficiency in the target language develops, qualitative patterns of sensitivity to linguistic and task features become more native-like and, furthermore, findings suggest that at lower levels of proficiency, learners strategically exploit their explicit linguistic knowledge—to varying degrees of success—to compensate for deficits in their implicit linguistic competence that is still in a state of flux and development.

Now that I have outlined in what specific ways this thesis research makes a novel contribution to the fields of late acquisition research and multilingualism, I would like to conclude by considering how the present findings may contribute to language teaching practice. What can these findings on the learner-external and learner-internal predictive factors of performance tell us about how language is acquired in the instructed setting and how language learning outcomes may be improved upon for adult learners? Recall that despite my best recruitment efforts from a variety of levels of Spanish coursework, the vast majority of adult learners in the present sample tested at the beginner level of proficiency. Therefore, there appears to be a disconnect between target proficiency level and actual proficiency level and, therefore, university-level language teaching could stand to benefit from a reconsideration of what factors may most contribute to language acquisition in adult learners and how these factors may be addressed and better supported in the classroom context.

First of all, the present findings suggest that adult learners are sensitive to certain aspects of linguistic stimuli, principally the relative frequency and morphology of nouns, and tend to assume that any given stimulus is grammatical unless known otherwise. These natural sensitivities could be exploited and supported through explicit instruction that highlights the canonical patterns of the correspondence between the derivational suffixes of nouns in Spanish and their gender assignment to more effectively and quickly direct learners' attention to these important and reoccurring patterns. This could be achieved through Processing Instruction (see VanPatten, 2014, 2015, 2018) with task design features that manipulate the stimuli to direct learners to focus on form by processing morpho-lexical units in the input, encouraging optimal form processing strategies while learners simultaneously interpret meaning to successfully complete a given task. For example, one could present learners with a forced-choice to complete an idea in which a given object is described in Spanish using an image that depicts an overtly-gendered noun and two corresponding overtly-gendered adjectives—one overtly marked as masculine, and one overtly marked as feminine—to choose from in order to describe the given object depicted in the picture. The gendered noun suffixes on the target noun and corresponding adjectives from which to choose could be highlighted to encourage learners to notice these morphological patterns. In addition to manipulating noun modifiers to highlight gendered morphology, learners' attention could also be directed to the corresponding agreement between the determiner phrase and the noun phrase using a similar processing instruction procedure in which learners are prompted with a gendered article and are asked to select the corresponding noun that would logically complete the phrase. It is essential here that learners are not just focusing on noun, article, and adjective morphological form, but that they are also processing the linguistic input for meaning, in order to encourage more intuitive and implicit language processing that will more substantially contribute to their developing (implicit) linguistic competence. Furthermore, given that learners are sensitive to the relative frequency of target nouns in their processing of grammatical gender agreement, these frequencies could be manipulated in the instructional material to ensure that particularly problematic nouns, such as those that are relatively infrequent and non-overtly marked for gender, are presented more frequently, to encourage the noticing (see Schmidt, 2001) of their corresponding gender assignment based on the linguistic context, that is, as learners interpret gender assignment on overtly gendered determiners and modifiers based on morphological cues. This instructional manipulation of the linguistic input to which learners are exposed would help to override the natural frequency effects of nouns whose gender assignment and corresponding agreement present particular difficulty for adult learners.

Second, we have observed that adult performance is modulated by the task demands and that adult learners struggle more with linguistic stimuli presented in the aural modality and with tasks under a time pressure. In other words, adults seem to naturally perform better with written stimuli and with tasks in which they can take their time, likely due to the fact that at lower levels of proficiency, learners' explicit linguistic knowledge in declarative memory is developed to a relatively larger extent than their implicit linguistic competence in procedural memory. Therefore, instruction for adult learners should integrate more tasks targeted at the speeded processing of auditory input. However, it is essential to consider what constitutes comprehensible input (see Krashen, 1985) at varying levels of proficiency so as not to overwhelm learners with tasks that are too challenging. Therefore, instructional tasks could take a given set of target stimuli, at the targeted proficiency level, and provide repeated input in both the auditory and written modalities followed by subsequent learning assessments that test both modalities together under both speeded and self-paced testing conditions. Furthermore, related to the previous discussion of sensitivity to noun frequencies, certain stimuli that are found via testing to be more difficult than others (i.e., produce relatively more errors) could be presented more often in subsequent input in both auditory and written modalities to further refine adult learners' sensitivity to linguistic features that present the most difficulty.

Third, the finding that the majority of adult learners report multilingual exposure and the observed multilingual learner advantage associated with prior language typology suggest that learners' other known languages could potentially be exploited in the classroom context to improve learning outcomes. Perhaps if learners are made more explicitly aware of how the linguistic features of their other known languages may relate to the target language being learned in the classroom, such as cross-linguistic congruency of gender assignment between Spanish and other known gendered languages, they could more effectively exploit this knowledge to their advantage. Furthermore, language teachers could request that their multilingual students verbalize other language knowledge that they have as it relates to the target structure and/or lesson at hand, as a sort of comparative grammar analysis to scaffold both their own learning and the learning of their classmates. Moreover, the findings of the present study suggest that a typological multilingual learner advantage is more associated with implicit linguistic competence; since metalinguistic awareness was also found to significantly and positively contribute to learner performance, leading multilingual learners to more explicitly reflect on their linguistic knowledge from their other languages known—particularly as it pertains to a structure being learned in class—could encourage the development of explicit awareness of language structure more globally, potentially increasing the scaffolding effect of metalinguistic awareness. Nonetheless, recall that no significant correlation was detected between prior linguistic repertoire and multilingualism in the present study, suggesting that these two factors might not necessarily be connected; however, explicit instructional focus on exploiting other language knowledge might facilitate the connection between multilingualism and metalinguistic awareness so that late learners may benefit not only from explicit awareness of the target language but also from the explicit awareness of language structure in the other languages in which they likely already have extensive implicit linguistic competence. Finally, connectionist theories of language acquisition suggest that connecting new knowledge to pre-existing forms of knowledge will strengthen the connection and facilitate longer-term learning. Therefore, encouraging multilingual students to make analytical comparisons between the target language in the classroom and their prior linguistic knowledge from other languages previously learned would help to facilitate and deepen this connection.

Fourth, the present findings suggest that metalinguistic awareness is highly predictive of learner performance and is particularly helpful in enhancing performance on cognitively offline tasks that allow for sufficient processing time to tap into explicit knowledge in declarative memory. Furthermore, it has been observed through regression modeling of significant predictors of learner performance at different proficiency levels that metalinguistic awareness is most predictive of learner performance at the beginner proficiency level and that all learners considered as a group outperformed their native speaker peers on the metalinguistic awareness survey. This suggests that learner performance is largely scaffolded by what learners know explicitly about the target structure and this knowledge can be strategically exploited when the task conditions permit. Instructional tasks in the classroom could first encourage the (natural) development of explicit awareness of language structure through the interpretation and analysis of many examples of a given grammar structure, such as gender agreement, and then once the explicit representation of this linguistic knowledge has been sufficiently rehearsed, instruction should push learners to operationalize this explicit knowledge about language form in real time under a time constraint and through responding to linguistic stimuli presented in the auditory modality. In other words, classroom instruction should work to move adult language learners from the explicit processing and analysis of linguistic information—which comes so naturally to them, particularly in the instructed context—to the implicit processing of auditory linguistic input and fluent real-time production under a time constraint.

Fifth, regarding the effect of motivation, the finding that the Ideal L2 Self component of motivation appears to be most strongly correlated with accuracy and is also more associated with online (time-constrained) and auditory language processing—which adult learners at beginner and intermediate levels find most challenging—suggests that it may be important to explicitly discuss with language learners the impact that a strong motivational orientation can have on their development in the target language. Perhaps encouraging learners to become more aware of their own individual motivations to learn the language could have a reinforcing effect on how motivated they feel and thereby how successful they may become in their acquisition of the target language. However, the exact connection between motivation and proficiency is less clear as one cannot be sure from the present dataset of the direction of influence between motivation and proficiency, that is, whether motivation leads to proficiency or whether proficiency may also lead to motivation. In any case, the present findings show that motivation to learn Spanish is positively correlated with weekly Spanish use, metacognitive awareness of one's own learning process, and explicit awareness of the target structure. Therefore, these data seem to suggest that motivation leads to positive cascading effects that may manifest differently in different learners and may differentially impact learning throughout the language acquisition process. Moreover, the present analysis has elucidated the nature of the motivation effect such that the Ideal L2 Self was identified as the subcomponent of motivation most positively correlated with linguistic performance, thereby demonstrating that the more one chooses to integrate their learning of the target language into their ideal self-identity, that is, how they would like to be perceived by those around them, the more a positive motivational orientation helped them in their performance with grammatical gender. Therefore, in the classroom setting, adult learners could be encouraged to outline the exact ways in which learning the target language will further contribute to the person that they would like to become; if they are able to clearly reflect on this and articulate this image of their ideal self-identity as someone who is learning a new language, that is, their Ideal L2 Self, the more they stand to benefit from the positive impact of motivation on their linguistic performance. In other words, language teachers must not only consider the linguistic and task factors that affect student learning of language in the classroom context, but also the individual characteristics of learners that can be encouraged through guided and informed reflection to potentially improve learning outcomes.

Finally, the present findings on how proficiency level modulates the relative effect that both linguistic and task factors have on adult learner performance suggest that although learners maintain qualitatively similar patterns of performance with regards to the features of linguistic stimuli throughout their late language acquisition trajectory, they nonetheless are able to more closely approximate native speaker norms of performance in a quantitative capacity as their proficiency develops, suggesting a different developmental trajectory that eventually results in indistinguishably similar observable linguistic behavior to that of native speakers, at least in the morphosyntactic domain of language structure. Furthermore, the finding that advanced learners patterned with native speakers in terms of their relative sensitivities to task effects, moving from enhanced accuracy on self-paced tasks presented in the written modality at beginner and intermediate proficiency levels to enhanced accuracy on online speeded tasks with no stimuli modality differences at the advanced proficiency level, suggests that late learner language performance is supported by explicit knowledge sources at the onset and intermediate stages of development, while at more advanced levels, linguistic behavior becomes gradually more supported by implicit and procedural knowledge. Therefore, language pedagogy should conceptualize the late learner developmental trajectory as one that would benefit from the scaffolding effect of explicit language knowledge and metalinguistic awareness of grammatical structure since language knowledge in adult learners gradually and naturally transitions from more explicit-based in declarative memory to more implicit-based in procedural memory. The present empirical findings seem to suggest that this transition from explicit to implicit language knowledge would likely be encouraged by a comparable transition in pedagogical task design between cognitively offline tasks in the written modality—allowing learners processing time to tap into and rehearse their explicit linguistic knowledge—to more cognitively online tasks in the auditory modality, requiring access to the developing implicit linguistic system in real-time under a time constraint. Adult learners will be better able to respond to varied task demands with greater levels of accuracy if given the opportunity to consolidate their explicit linguistic knowledge through continued practice with tasks that demand the online use and auditory processing of this knowledge, thereby contributing more substantially to the underlying development of their implicit linguistic competence. Often times we unfairly expect learners to take their explicit knowledge about language structure learned in the classroom environment and apply it in real-time to communicative situations in the real world. This abrupt disconnect between pedagogy and real-world linguistic expectations is both frustrating and de-motivating for students and could be mitigated through the application of varied and strategic task design features that integrate offline/online processing and written/auditory modalities at both the instructional and assessment phases.

In sum, it is essential that empirical findings that elucidate how adult learners acquire a new language and the factors that most contribute to their performance do not stay merely within the language acquisition research community, but rather be transmitted more broadly and translated into pedagogical concepts. Empirical findings on the impact of individual learner characteristics and the relative sensitivities that adult learners naturally exhibit to linguistic stimuli and language task features should be considered when designing pedagogical materials and when deciding on what language material to teach and *how* to teach it—that is, more implicitly or more explicitly by conditioning online or offline processing—and at what proficiency level. We know that adult learners exhibit extensive variation in their performance and ultimate attainment, and I have shown here through this thesis research that much of this variation can be explained in terms of sensitivities to linguistic features and task constraints as well as by individual learner characteristics. I argue that adult learners only stand to benefit from a language pedagogy that strategically considers the growing body of late language acquisition research when making important pedagogical decisions regarding language curricula, learning activities, and assessments. Such a research-based language pedagogy would crucially consider adult learners as fully capable of acquiring language competence that is outwardly indistinguishable from native speaker linguistic behavior—at least in the domain of morphosyntax—while also recognizing that the developmental trajectory to arrive at such competence likely differs in important ways from native speakers and early learners and therefore should be supported differently too. We have observed how late learners in the instructed context from beginner to advanced levels of proficiency are effective at developing explicit linguistic knowledge through the direct analysis of the target structures being taught. Adult learners in the classroom context are clearly capable of forming both explicit and implicit representations of linguistic information and their global proficiency level appears to be largely dependent upon the relative balance between these two sources of language knowledge. Therefore, if as language teachers we can support learners' natural tendencies to be more accurate with frequent linguistic items and written and self-paced language processing while also encouraging a gradual transition to more online and implicit processing through strategic task manipulation in both language teaching and language assessment, we can more effectively move adult learners along their natural explicit-to-implicit nonnative developmental trajectory to improve language learning outcomes. Despite the immense variation in ultimate attainment, I assert that adults can fully acquire additional languages with the proper pedagogical support that strategically considers and optimizes their natural language processing strategies and biases and that acknowledges both the importance and potential of individual differences.

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language learning*, *59*(2), 249-306.

Ajzen, I. (1988). *Attitudes, Personality and Behavior*. Chicago: Dorsey Press.

Alarcón, I. V. (2011). Spanish gender agreement under complete and incomplete acquisition: Early and late bilinguals' linguistic behavior within the noun phrase. *Bilingualism: Language and Cognition*, 14(3), 332-350.

Anderson, J.R. (1995). *Learning and Memory: An Integrated Approach*. New York: Wiley.

Andringa, S. J. (2014). The use of native speaker norms in critical period hypothesis research. *Studies in Second Language Acquisition*, *36*, 565–596. https://doi.org/10.1017/S0272263113000600

Andringa, S., & Dąbrowska, E. (2019). Individual differences in first and second language ultimate attainment and their causes. *Language Learning*, *69*(S1), 5-12.

Au, S.Y., (1988). A critical appraisal of Gardner's socio-psychological theory of second-language (L2) Learning. *Language Learning*, 38, 75–100.

Aubrey, S., & Nowlan, A. G. (2013). Effect of intercultural contact on L2 motivation: A comparative study. *Language learning motivation in Japan*, *71*, 129.

Bandura, A., (1997). *Self-efficacy: The Exercise of Control*. New York: Freeman.

Beatty-Martínez, A. L., & Dussias, P. E. (2019). Revisiting masculine and feminine grammatical gender in Spanish: linguistic, psycholinguistic, and neurolinguistic evidence. *Frontiers in psychology*, *10*, 751.

Black, M. & Tararova, O. (2020). Adult Acquisition of Grammatical Gender in Instructed L2 Spanish and the Role of Metacognition. Proceedings of the 2020 annual conference of the Canadian Linguistic Association, online, May 30-June 1, 2020.

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language learning*, *29*(1), 81-103.

Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Applied Linguistics*, *3*(3), 181-206.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in second language acquisition*, *16*(2), 157-168.

Bialystok, E. (1997). Effects of bilingualism and biliteracy on children's emerging concepts of print. *Developmental psychology*, *33*(3), 429.

Birdsong, D. (2005). Nativelikeness and non-nativelikeness in L2A research. *International Review of Applied Linguistics*, *43*, 319–328. https://doi.org/10.1515/iral.2005.43.4.319

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, *56*, 9–49.

Birdsong, D., Gertken, L.M., & Amengual, M. (2012). *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism.* COERLL, University of Texas at Austin. Web. 20 Jan. 2012. https://sites.la.utexas.edu/bilingual/

Birdsong, D., & Vanhove, J. (2016). Age of second language acquisition: Critical periods and social concerns. In S. Nicoladis & E. Montanari (Eds.), *Bilingualism across the lifespan: Factors moderating language proficiency* (pp. 163–181). Washington, DC: American Psychological Association. https://doi.org/10.1037/14939-010

Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge: What can heritage language learners contribute?. *Studies in second language acquisition*, *33*(2), 247-271.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory & Cognition*, *41*(2), 281–296.

Brooks, P. J., Kwoka, N., & Kempe, V. (2017). Distributional effects and individual differences in L2 morphology learning. *Language Learning*, *67*(1), 171-207.

Bruhn de Garavito, J. B. (2007). Acquisition of the Spanish plural by French L1 speakers: the role of transfer. *The role of formal features in second language acquisition*, 271-298.

Bruhn de Garavito, J. L. (1999). The syntax of Spanish multifunctional clitics and near-native competence. *Applied Linguistics Review*, 3(2), 247-295.

Bruhn de Garavito, J. B., & Valenzuela, E. (2008). Eventive and stative passives in Spanish L2 acquisition: A matter of aspect. *Bilingualism: Language and Cognition*, *11*(3), 323-336.

Bruhn de Garavito, J., & White, L. (2002). The second language acquisition of Spanish DPs: The status of grammatical features. In A. T. Pérez-Leroux & J. M. Liceras (eds.), *The acquisition of Spanish morphosyntax: The L1/L2 connection*, pp. 153–178. Dordrecht: Kluwer.

Bybee, J. L., & Hopper, P. J. (Eds.). (2001). *Frequency and the emergence of linguistic structure* (45). John Benjamins Publishing.

Caplan, D., Alpert, N., & Waters, G. (1999). PET studies of syntactic processing with auditory sentence presentation. *NeuroImage*, *9*(3), 343-351.

Carpenter, H. S. (2008). *A behavioral and electrophysiological investigation of different aptitudes for L2 grammar in learners equated for proficiency level* (Doctoral dissertation, Georgetown University).

Cenoz, J. (2013). The influence of bilingualism on third language acquisition: Focus on multilingualism. *Language teaching*, *46*(1), 71.

Chen, L., & Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Attention, Perception, & Psychophysics*, *75*(5), 790-811.

Cho, M. (2018). Task complexity and modality: Exploring learners' experience from the perspective of flow. *Modern Language Journal, 102*, 162–180.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching*, 5, 161-170.

Crookes, G., & Schmidt, R. (1991). Motivation: Reopening the research agenda. *Language Learning*, 41, 469–512.

Crowder, R. G. (1986). Auditory and temporal factors in the modality effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *2*, 268–278.

Cummins, J. (2007). Rethinking monolingual instructional strategies in multilingual classrooms. *Canadian Journal of Applied Linguistics* 10(2), 221–240.

Cummins, J. (2009). Multilingualism in the English-language Classroom: Pedagogical Considerations. *TESOL Quarterly* 43(2). 317–321. https://doi.org/10.1002/j.1545-7249.2009.tb00171.x

Dąbrowska, E. (2012). Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism*, *2*, 219–253. http://doi.org/10.1075/lab.2.3.01dab

Dąbrowska, E. (2019). Experience, Aptitude, and Individual Differences in Linguistic

Attainment: A Comparison of Native and Nonnative Speakers. *Language Learning*, *69*(S1), 72 100. https://doi.org/10.1111/lang.12323

Dąbrowska, E., & Street, J. (2006). Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. *Language Sciences*, *28*(6), 604-615.

Deci, L., Ryan, M., (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. New York: Plenum.

DeKeyser, R. (1998). Beyond focus on form. *Focus on form in classroom second language acquisition*, 42-63.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in second language acquisition*, *22*(4), 499-533.

DeKeyser, R. (2012). Interactions between individual differences, treatments, and structures in SLA. *Language Learning*, *62,* 189-200.

Dewaele, J.M., K.V. Petrides, & A. Furnham. (2008). Effects of Trait Emotional Intelligence and Sociobiographical Variables on Communicative Anxiety and Foreign Language Anxiety Among Adult Multilinguals: A Review and Empirical Investigation. *Language Learning* 58. 911–960. https://doi.org/10.1111/j.1467-9922.2008.00482.x

Dörnyei, Z. (2000). Motivation in action: Towards a process‐oriented conceptualisation of student motivation. *British journal of educational psychology*, *70*(4), 519-538.

Dörnyei, Z. (2001). New themes and approaches in second language motivation research. *Annual review of applied linguistics*, *21*(1), 43.

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dörnyei, Z. (2009). The L2 Motivational Self System. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Bristol, UK: Multilingual Matters.

Dörnyei, Z., & Csizér, K. (2002). Some dynamics of language attitudes and motivation: Results of a longitudinal nationwide survey. *Applied Linguistics*, *23*, 421–462.

Dörnyei, Z., & Ottó, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in Applied Linguistics* (Thames Valley University, London), 4, 43–69.

Dörnyei, Z., & Skehan, P. (2008). Individual Differences in Second Language Learning. In *The Handbook of Second Language Acquisition* (pp. 589–630). John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470756492.ch18

Duffield, N., & White, L. (1999). Assessing L2 knowledge of Spanish clitic placement: Converging methodologies. *Second Language Research*, *15*(2), 133-160.

Eagly, A. H. & Chaiken, S. (1993). *The Psychology of Attitudes*. New York: Harcourt Brace.

Eddington, D. (2002). Spanish gender assignment in an analogical framework. *Journal of Quantitative Linguistics*, *9*(1), 49-75.

Eddington, D., & Hualde, J. I. (2008). El abundante agua fría: Hermaphroditic Spanish Nouns. *Studies in Hispanic & Lusophone Linguistics*, *1*(1).

Eichenbaum, H. (2002). *The cognitive neuroscience of memory: An introduction*. New York: Oxford University Press.

Eichenbaum, H., & Cohen, N. (2001). From conditioning to conscious recollection: Memory systems of the brain. New York: Oxford University Press.

Ellis, N. C. (2015). Implicit and explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Studies in Bilingualism* (Vol. 48, pp. 1–24). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.48.01ell

Ellis, N. C. (2007). Dynamic systems and SLA: The wood and the trees. *Bilingualism: Language and cognition*, *10*(1), 23-25.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in second language acquisition*, 305-352.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in second language acquisition*, *24*(2), 143-188.

Ellis, N. C. (1994). *Vocabulary acquisition: The implicit ins and outs of explicit cognitive mediation*. Academic Press.

Ellis, N. C. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *The Quarterly Journal of Experimental Psychology Section A*, *49*(1), 234-250.

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style in the use of the past tense. *Studies in Second Language Acquisition*, 9: 12-20.

Ellis, R. (1989). *Understanding second language acquisition* (Vol. 31). Oxford: Oxford university press.

Ellis, R. (1993). The structural syllabus and second language acquisition. *TESOL quarterly*, *27*(1), 91-113.

Ellis, R. (2005). Measuring Implicit and Explicit Knowledge of a Second Language: A Psychometric Study. *Studies in Second Language Acquisition*, *27*(02). https://doi.org/10.1017/S0272263105050096

Ellis, N. C., & Larsen-Freeman, D. (2009). *Language as a complex adaptive system* (Vol. 11). John Wiley & Sons.

Ellis, N. C. & D. Larsen-Freeman (2006). Language emergence: Implications for applied linguistics. Introduction to the special issue. *Applied Linguistics* 27.4, 558–589.

Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in second language acquisition*, *19*(2), 145-171.

Engle, R. W., & Mobley, L. A. (1976). The modality effect: What happens in long-term memory?. *Journal of Verbal Learning and Verbal Behavior*, *15*(5), 519-527.

Farmer, T. A., Misyak, J. B., & Christiansen, M. H. (2012). Individual differences in sentence processing. *Cambridge handbook of psycholinguistics*, *353*, 364.

Fernández-García, M. (1999). Patterns of gender agreement in the speech of second language learners. In *Advances in Hispanic linguistics: Papers from the 2nd Hispanic linguistics symposium* (pp. 3-15). Somerville, MA: Cascadilla Press.

Foote, R. (2015). The production of gender agreement in native and L2 Spanish: The role of morphophonological form. *Second Language Research*, *31*(3), 343–373.

Foster, P., & Tavakoli, P. (2009). Native speakers and task performance: Comparing effects on complexity, fluency, and lexical diversity. *Language learning*, *59*(4), 866-896.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*, 117–125.

Gamboa Rengifo, A. (2012). *The acquisition of grammatical gender in Spanish by English-speaking L2 learners* (M.A. thesis), Purdue University. https://search.proquest.com/docview/1239142014/abstract/8F305596B604699PQ/1

García Magdali, L. (2010). Metacognitive strategies-based instruction to support learner autonomy in language learning. *Revista Canaria de Estudios Ingleses*, *61*, 73–86.

Gardner, R. C. (1985). *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation.* London: Edward Arnold.

Gardner, R. C. & Lambert, W. E. (1972). *Attitudes and Motivation in Second Language Learning*. Rowley, MA: Newbury House.

Gardner, R. C., & Lysynchuk, L. M. (1990). The role of aptitude, attitudes, motivation, and language use on second-language acquisition and retention. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *22*(3), 254.

Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, *29*, 311–343. https://doi.org/10.1177/0267658312461497

Gregg, K. R. (1989). Second language acquisition theory: The case for a generative perspective. *Linguistic perspectives on second language acquisition*, 15-40.

Gries, S. T., & Divjak, D. (Eds.). (2012). *Frequency effects in language learning and processing* (Vol. 244). Walter de Gruyter.

Grüter, T., Lew-Williams, C., & Fernald, A. (2012). Grammatical gender in L2: A production or a real-time processing problem? *Second Language Research*, *28*(2), 191-215.

Halliday, M. A. K. (2007). On the concept of 'educational linguistics'. In J. Webster (ed.), *The collected works of M. A. K. Halliday, Volume 9: Language and education*. London: Continuum, 354–367. Originally published in R. Giblett & J. O'Carroll (1990) (eds.), *Discipline, dialogue, difference: Proceedings of the Language in Education conference, Murdoch University, December 1989.* Murdoch, Australia: Duration, 23–42.

Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, *44*, 9–15. https://doi.org/10.1016/j.lindif.2015.10.003

Harris, J. W. (1991). The exponence of gender in Spanish. *Linguistic Inquiry*, 22(1), 27-62.

Hernández Pina, F. (1984). *Teorías psicosociolingüísticas y su aplicación a la adquisición del español como lengua materna.* Madrid: Siglo XXI.

Hirosh, Z., & Degani, T. (2018). Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review*, *25*(3), 892-916.

Hudson, R. A. (2007). English dialect syntax in Word Grammar. *English Language & Linguistics*, *11*(2), 383-405.

Hudson, R. (2021). Word grammar. In *The Routledge Handbook of Cognitive Linguistics* (pp. 111-126). Routledge.

Hulstijn, J.H. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, *18*, 193–223.

Hulstijn, J.H. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning: The virtues and pitfalls of a two-system view. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages*. Amsterdam, Netherlands: John Benjamins. https://doi.org/10.1075/sibil.48.02hul

Hulstijn, J. H. (2019). An individual-differences framework for comparing nonnative with native speakers: Perspectives from BLC theory. *Language Learning*, *69*, 157-183.

Hult, F. (2010). Theme-based research in the transdisciplinary field of educational linguistics. In F. Hult (ed.), *Directions and prospects for educational linguistics*. Dordrecht: Springer, 28–50.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 538-588). Malden, MA: Blackwell Pub.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: individual differences in working memory. *Psychological review*, *99*(1), 122.

Kaushanskaya, M. & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review* 16. 705–710. https://doi.org/10.3758/PBR.16.4.705

Khodabakhshzadeh, H., Hosseinnia, M., & Rahimian, S. (2017). Learning style, metacognition and creativity as predictors of the foreign language achievement: A structural equation modeling approach. *Psychological Studies*, *62*(4), 377–385.

Kissling, E. M., & O’Donnell, M. E. (2015). Increasing language awareness and self-efficacy of FL students using self-assessment and the ACTFL proficiency guidelines. *Language Awareness*, *24*(4), 283–302.

Klein, E. C. (1995). Second versus Third Language Acquisition: Is There a Difference? *Language Learning*, *45*(3), 419–466. https://doi.org/10.1111/j.1467-1770.1995.tb00448.x

Knowlton, B., & Moody, T. (2008). Procedural learning in humans. In J. Byrne (Ed.), Learning and memory: A comprehensive reference. Memory systems, vol 3. (pp. 321–340). Oxford: Elsevier.

Kohler, D. B. (2002). The effects of metacognitive language learning strategy training on lower-achieving second language learners (Doctoral dissertation). Brigham Young University. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 63(5-A), 1690.

Kormos, J. (2006). The structure of self-repairs in the speech of Hungarian learners of English. *Acta Linguistica Hungarica*, *53*(1), 53-76.

Kormos, J. (2011). Task complexity and linguistic and discourse features of narrative writing performance. *Journal of Second Language Writing*, *20*(2), 148-161.

Kormos, J. (2014). Differences across modalities of performance: An investigation of linguistic and discourse complexity in narrative tasks. In H. Byrnes & R. M. Manchón (Eds.), *Task-based language learning—Insights from and for L2 writing* (pp. 193–217). Philadelphia/Amsterdam: John Benjamins.

Krashen, S. (1977). Some issues relating to the monitor model. *On Tesol*, *77*(144-158).

Krashen, S. D. (1981). Bilingual education and second language acquisition theory. *Schooling and language minority students: A theoretical framework*, 51-79.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. Lincolnwood, IL: Laredo.

Kupisch, T., & Rothman, J. (2018). Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22(5), 564-582. doi:10.1177/1367006916654355

Larsen-Freeman, D. (1978). An ESL index of development. *TESOL quarterly*, 439-448.

Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied linguistics*, *18*(2), 141-165.

Larsen-Freeman, D. (2002). Making sense of frequency. *Studies in second language acquisition*, *24*(2), 275-285.

Larsen-Freeman, D. (2006). The Emergence of Complexity, Fluency, and Accuracy in the

Oral and Written Production of Five Chinese Learners of English. *Applied Linguistics*, *27*(4), 590–619. https://doi.org/10.1093/applin/aml029

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied linguistics*, *30*(4), 579-589.

Larsen-Freeman, D. (2012). Complex, dynamic systems: A new transdisciplinary theme for applied linguistics? *Language Teaching*, *45*(2), 202–214. https://doi.org/10.1017/S0261444811000061

Larsen-Freeman, D., & Cameron, L. (2008). Research Methodology on Language Development from a Complex Systems Perspective. *The Modern Language Journal*, *92*(2), 200–213. https://doi.org/10.1111/j.1540-4781.2008.00714.x

Leaver, B.L., (2003). Motivation at native-like levels of foreign language proficiency: a research agenda. *Journal for Distinguished Language Studies*, 1 (1).

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*(3), 193-198.

Lightbown, P. M., Spada, N., & White, L. (1993). The role of instruction in second language acquisition. *Studies in Second Language Acquisition, 15 [Special issue]*.

López Prego, B. (2015). The online use of markedness information in L1 and L2 Spanish gender agreement. Doctoral dissertation. Lawrence, Kansas: University of Kansas, USA.

Markus, H., & Nurius, P. (1986). Possible selves. *American psychologist*, *41*(9), 954.

Markus, H., & Nurius, P. (1987). Possible selves. *American Psychologist*, 41, 954–969.

May, S. (2014). Introducing the “multilingual turn”. *The multilingual turn: Implications for SLA, TESOL and bilingual education*, 1-6.

McCarthy, C. (2008). Morphological variability in the comprehension of agreement: An argument for representation over computation. *Second Language Research*, *24*(4), 459–486. https://doi.org/10.1177/0267658308095737

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in second language acquisition*, (20): 52-83.

Michel, M. C., Révész, A., Shi, D., & Li, Y. (2019). Chapter 6. The effects of task demands on linguistic complexity and accuracy across task types and L1/L2 speakers. In Z. (Edward) Wen & M. J. Ahmadian (Eds.), *Task-Based Language Teaching* (Vol. 13, pp. 133–151). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.13.07mic

Montrul, S. (2002). Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and cognition*, *5*(1), 39-68.

Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Montrul, S., Foote, R., & Perpiñán, S. (2008). Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, *58*(3), 503–553.

Morgan-Short, K., & Ullman, M.T. (2012). The neurocognition of second language. In S. Gass, & A. Mackey (Eds.), *The Routledge handbook of second language acquisition*: 282–299. New York: Routledge.

Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K.A., Carpenter, H., &Wong, P.C.M. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism: Language and Cognition*, 17, 56–72.

Morgan-Short, K., Deng, Z., Brill-Schuetz, K.A., Faretta-Stutenberg, M., Wong, P.C.M., & Wong, F.C.K. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Studies in Second Language Acquisition*, 37: 383–419.

Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. (2012). Second language processing shows increased native-like neural responses after months of no exposure. *PloS One*, 7, 1–18.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24, 933–947.

Murphy, V. A. (1997). The effect of modality on a grammaticality judgement task. *Second Language Research*, *13*(1), 34–65. https://doi.org/10.1191/026765897671676818

Newell, A. & H. Simon (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice Hall.

Nosratinia, M., Saveiy, M., & Zaker, A. (2014). EFL learners' self-efficacy, metacognitive awareness, and use of language learning strategies: how are they associated? *Theory and Practice in Language Studies*, *4*(5), 1080–1092.

Ochs, E. (1979). Planned and unplanned discourse, in T. Givon (ed.): *Syntax and Semantics, Vol. 12: Discourse and Semantics*. New York: Academic Press.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21: 109-148.

Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *The modern language journal*, *78*(1), 12-28.

Oxford, R. L. (1996). Employing a questionnaire to assess the use of language learning strategies. *Applied language learning*, *7*(1), 28-47.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*(12), 2728-2744.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: implication for bilingualism and second language acquisition. *Implicit and explicit learning of languages*.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism* (Vol. 18). Philadelphia: John Benjamins Publishing.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Philadelphia: John Benjamins Pub.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory & cognition*, *17*(4), 398-422.

Pérez Pereira, M. (1991). The acquisition of gender: What Spanish children tell us. *Journal of Child Language* 18(3). 571–590.

Perdue, C. (Ed.). (1993). *Adult language acquisition: Crosslinguistic perspectives*. Cambridge: Cambridge University Press.

Prévost, P., & White, L. (2000). Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, *16*(2), 103–133. https://doi.org/10.1191/026765800677556046

Recanzone, G. H. (2009). Interactions of auditory and visual stimuli in space and time. *Hearing Research*, *258*, 89–99.

Rothman, J. (2013). Cognitive Economy, Non-Redundancy and Typological Primacy in L3 Acquisition: Evidence from initial stages of L3 Romance. In S. Baauw, F. Drijkoningen, L. Meroni, & M. Pinto (Eds.), *Romance Languages and Linguistic Theory 2011* (pp. 217-247). Amsterdam/Philadelphia: John Benjamins.

Rothman, J. (2014). Linguistic and cognitive motivations for the typological primacy model of third language (L3) transfer: considering the role of timing of acquisition and proficiency in previous languages. *Bilingualism: Language and Cognition*.

Rothman, J., Cabrelli Amaro, J., & de Bot, K. (2013). Third language acquisition. In J. Herschensohn & M. Young Scholten (Eds.), *Cambridge handbook of second language acquisition* (pp. 372-393). Cambridge: Cambridge University Press.

Saeedi, Masoud. (2020). Task Condition and L2 oral Performance: Investigating the Combined Effects of Online Planning and Immediacy. *International Journal of Foreign Language Teaching & Research*, *8*(32), 35–48.

Sagarra, N., & Herschensohn, J. (2010). The role of proficiency and working memory in gender and number agreement processing in L1 and L2 Spanish. *Lingua*, *120*(8), 2022–2039. https://doi.org/10.1016/j.lingua.2010.02.004

Sagarra, N., & Herschensohn, J. (2013). Processing of gender and number agreement in late Spanish bilinguals. *International Journal of Bilingualism*, *17*(5), 607-627.

Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied linguistics*, *11*(2), 129-158.

Schmidt, R. (2001). Attention. In P. Robinson (Ed), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. *Talking to learn: Conversation in second language acquisition*, *237*, 326.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, *19*(4), 460–475.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and learning*, *4*, 33-45.

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in second language acquisition*, 15(2), 147-163.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209–31.

Singleton, D., & Muñoz, C. (2011). Around and beyond the critical period hypothesis. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. *II*; pp. 407–425). London, England: Routledge.

Smith, M. S. (1981). Consciousness-raising and the second language Learner1. *Applied linguistics*, *2*(2), 159-168.

Skehan, P. (1996). Second-language acquisition research and task-based instruction, in J. Willis and D. Willis (Eds.): *Challenge and Change in Language Teaching*. Oxford: Heinemann, 17-30.

Skehan, P., & Foster, P. (1997). Task type and task processing conditions as influences on foreign language performance. *Language teaching research*, *1*(3), 185-211.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, *30*(4), 510-532.

Skehan, P., & Shum, S. (2014). Structure and processing condition in video-based narrative retelling. *Processing perspectives on task performance*: 187-210.

Skehan, P. (2015). Foreign language aptitude and its relationship with grammar: A critical overview. *Applied Linguistics*, *36*(3), 367-384.

Spino-Seijas, A.L. (2017). Grammatical gender agreement in L2 Spanish: The role of syntactic context. Doctoral dissertation. East Lansing, Michigan: Michigan State University, USA.

Tagarelli, K.M. (2014). The neurocognition of adult second language learning: An fMRI study. (Doctoral dissertation) Department of Linguistics, Georgetown University.

Tanner, D., & Van Hell, J. G. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, *56*, 289-301.

Tavakoli, P., & Foster, P. (2008). Task design and L2 performance. *Language Learning*, *58*(2), 429-473.

Teschner, R. V., & Russell, W. M. (1984). The gender patterns of Spanish nouns: An inverse dictionary-based analysis. *Hispanic Linguistics*, *1*(1), 115-132.

Thompson, A.S. (2013). The interface of language aptitude and multilingualism: Reconsidering the bilingual/multilingual dichotomy. *The Modern Language Journal* 9 7(3). 685–701. https://doi.org/10.1111/j.1540-4781.2013.12034.x

Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*(2), 105–122. https://doi.org/10.1017/S1366728901000220

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M.T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition* (pp. 141–178). Washington, DC: Georgetown University Press.

Ullman, M.T. (2015). The declarative/procedural model: A neurobiologically motivated theory of first and second language. In B. VanPatten, & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 135–158) (2nd ed.). New York: Routledge.

Vandergrift, L., Goh, C. C., Mareschal, C. J., & Tafaghodtari, M. H. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language learning*, *56*(3), 431-462.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language Learning*, *60*(2), 470–497.

VanPatten, B. (1990). Attending to form and content in the input: An experiment in consciousness. *Studies in second language acquisition*, *12*(3), 287-301.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Greenwood Publishing Group.

VanPatten, B. (2014). *Input Processing in Adult SLA. In Theories in Second Language Acquisition* (2e éd.). Routledge.

VanPatten, B. (2015). Foundations of processing instruction. *International Review of Applied Linguistics in Language Teaching*, IRAL, 53(2), 91-109. https://doi.org/10.1515/iral-2015-0005

VanPatten, B. (2018). Processing Instruction. In *The TESOL Encyclopedia of English Language Teaching* (p. 1-7). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781118784235.eelt0094

Vasylets, O., Gilabert, R., & Manchon, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, *67*(2), 394-430.

Victori, M., & Lockhart, W. (1995). Enhancing metacognition in self-directed language learning. *System*, *23*(2), 223–234.

Walqui, A., (2000). Contextual factors in second language acquisition. ERIC Digest. ERIC Clearinghouse on Languages and Linguistics, Document ED444381, Washington, DC.

Wang, Z. (2014). Online time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27-62). Amsterdam, The Netherlands: John Benjamins.

Wang, Z., & Skehan, P. (2014). Structure, lexis, and time perspective. *Processing perspectives on task performance*, *5*, 155-185.

White, L. (1989). *Universal Grammar and Second Language Acquisition*. Amsterdam: John Benjamins.

White, L. (2003). *Second Language Acquisition and Universal Grammar*. Cambridge, UK: Cambridge University Press.

White, L., Valenzuela, E., Kozlowska–Macgregor, M., & Leung, Y. K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied psycholinguistics*, 25(1), 105-133.

Yuan, F., & Ellis, R. (2003). The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. *Applied Linguistics*, *24*(1), 1–27.

Zhang, R. (2015). Measuring University-Level L2 Learners' Implicit and Explicit Linguistic Knowledge. *Studies in Second Language Acquisition*, *37*(3), 457.

Appendices

Appendix A: Language Learner Profile Questionnaire

Language Learner Profile Questionnaire

→ Please enter the participant code you received: __________

A. Personal Information

1. Gender identification:
□ Female □ Male □ gender neutral/non-binary
□ You do not have an option that applies to me; I identify as ______________

2. Age (yrs.): ______

3. Occupation: ___________________________

4. Highest level of formal education *completed*:
□ less than high school □ high school □ CEGEP/College/Professional
□ some university □ university (B.A., B.S.) □ some graduate school
□ Master's □ PhD/MD/JD

B. First Language

5. What is your first language? ___________________________________

6. What is the first language of your…

mother? _______________________ *father?* _______________________

7. Did you learn your first language from birth? □ YES □ NO

If you answered 'No' to the question above, please explain:

____

8. Which language(s) did you speak at home as a child? (*list all that apply*)

____

9. Is your first language the language with which you are the most comfortable? □ YES □ NO

If you answered 'No' to the question above, please explain:
____

C. Language History

10. What is your birthplace (city, country)? ___________________________________

11. At what age (in years) did you start learning the following languages?

English: ____
Spanish: ____
Other (please specify): ______________
Other (please specify): ______________

12. At what age (in years) did you start to feel comfortable using the following languages?

English: ____
Spanish: ____
Other (please specify): ______________
Other (please specify): ______________

13. How many years of classes (grammar, history, math, etc.) have you had in the following languages (primary school through university)?

English: ____
Spanish: ____
Other (please specify): ______________
Other (please specify): ______________

14. How many years have you spent in a country/region where the following languages are spoken?

English: ____
Spanish: ____
Other (please specify): ______________
Other (please specify): ______________

15. How many years have you spent in a family where the following languages are spoken?

English: ____
Spanish: ____
Other (please specify): ______________
Other (please specify): ______________

16. How many years have you spent in a work environment where the following languages are spoken?

English: ____
Spanish: ____
Other (please specify): ______________
Other (please specify): ______________

D. Language Use

<u>Instructions</u>: Think about how you've used the languages you know during a typical week in the past month. Estimate what percentage of the time (0% → 100%) you use those languages in each of the following contexts. The total use for all languages in each context should equal 100%.

17. In an average week, what percentage of the time do you use the following languages with friends?

English: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Spanish: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%

18. In an average week, what percentage of the time do you use the following languages with family?

English: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Spanish: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%

19. In an average week, what percentage of the time do you use the following languages at school/work?

English: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Spanish: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%

20. When you talk to yourself, how often do you talk to yourself in the following languages?

English: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Spanish: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%

21. When you count, how often do you count in the following languages?

English: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Spanish: □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
Other: (specify): □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%

E. Language Proficiency: Self-assessment

Instructions: Please rate your current linguistic ability in *each* of your known languages in the following areas by clicking on the appropriate proficiency rating (Beginner, Intermediate, Advanced, Native-like).

SKILL	Beginner	Intermediate	Advanced	Native-like
22. READING				
English				
Spanish				
Other (specify)				
Other (specify)				
23. WRITING				
English				
Spanish				
Other (specify)				
Other (specify)				
24. SPEAKING				
English				
Spanish				

Other (specify)				
Other (specify)				
25. LISTENING				
English				
Spanish				
Other (specify)				
Other (specify)				
26. TOTAL COMPETENCE				
English				
Spanish				
Other (specify)				
Other (specify)				

F. Language Attitudes

<u>Instructions</u>: Think about how the languages you speak reflect your identity, values, and how you want to be perceived in society. There are no right or wrong answers.

27.

a. I feel like myself when I speak English.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

b. I feel like myself when I speak Spanish.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

c. I feel like myself when I speak ___________ (other language, *if applicable*).
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

d. I feel like myself when I speak ___________ (other language, *if applicable*).
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

28.

a. I identify with an English-speaking culture.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

b. I identify with a Spanish-speaking culture.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

c. I identify with a _________ (other language, *if applicable*)-speaking culture.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

d. I identify with a _________ (other language, *if applicable*)-speaking culture.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

29.

a. It is important to me to use English like a native speaker.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

b. It is important to me to use Spanish like a native speaker.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

c. It is important to me to use _________ (other language, *if applicable*) like a native speaker.
Strongly disagree > > > > > *neither/neutral* > > > > > > *Strongly agree*
□ 1 □ 2 □ 3 □ 4 □ 5

d. It is important to me to use _________ (other language, *if applicable*) like a native speaker.
Strongly disagree > > > > > neither/neutral > > > > > > Strongly agree
□ 1 □ 2 □ 3 □ 4 □ 5

30.

a. I want others to think I'm a native speaker of English.
Strongly disagree > > > > > neither/neutral > > > > > > Strongly agree
□ 1 □ 2 □ 3 □ 4 □ 5

b. I want others to think I'm a native speaker of Spanish.
Strongly disagree > > > > > neither/neutral > > > > > > Strongly agree
□ 1 □ 2 □ 3 □ 4 □ 5

c. I want others to think I'm a native speaker of ________ (other language, *if applicable*).
Strongly disagree > > > > > neither/neutral > > > > > > Strongly agree
□ 1 □ 2 □ 3 □ 4 □ 5

d. I want others to think I'm a native speaker of ________ (other language, *if applicable*).
Strongly disagree > > > > > neither/neutral > > > > > > Strongly agree
□ 1 □ 2 □ 3 □ 4 □ 5

Interim question:

Are you a native speaker of Spanish? □ YES □ NO

→ If 'YES', survey stops here.

Thank you for your responses to the Language Learner Profile Questionnaire! You may now proceed to your scheduled Zoom session with the researcher for the experimental portion of the study. If you have any questions or concerns at any time, please contact the researcher.
Thank you again and we look forward to meeting with you soon!

→ If 'NO', survey continues.

G. Language Learner Awareness

Instructions: Think of yourself as a student of Spanish and a language learner. Read each statement carefully. Consider how much each statement applies to you (in class, while studying, reviewing, practicing, etc.) in a typical week during the past month (or whenever you were last enrolled in a Spanish course). There are no correct answers; only answers that best represent you.
**Note:* if you are a native Spanish speaker, this section does not apply to you.

31. I ask myself if I learned as much as I could have once I finish a task or activity in Spanish.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

32. I know what information is most important to learn.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

33. I create my own examples in Spanish to make new information more meaningful.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

34. When I don't understand something in Spanish, I ask others for help (teacher, classmates, tutors, etc.).

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

35. I focus on overall meaning rather than specifics.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

36. I use different learning strategies depending on the situation.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

37. I change strategies when I fail to understand something.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

38. I stop and reread or listen again when I get confused by something in Spanish.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

39. I find myself using helpful learning strategies automatically.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

40. I ask myself if I am meeting my language learning goals.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

41. I understand my current strengths and weaknesses in Spanish.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

42. I am good at organizing new information that I learn in Spanish.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

43. I am aware of what strategies I use when I study.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

44. I am good at remembering new words and grammar concepts.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

45. I use the organizational structure of a text or activity to help me understand and learn from it.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

46. When listening to and reading Spanish, I consciously focus my attention on the most important information.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

47. I learn more Spanish when I am interested in the particular topic.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

48. When learning new information in Spanish (grammar and vocabulary), I try to use strategies that have worked in the past.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

49. Overall, I have control over how well I learn Spanish.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

50. I find myself pausing regularly to check my comprehension.

(1)	(2)	(3)	(4)	(5)
Never	Occasionally	Sometimes	Often	Always

H. Motivational Orientation

<u>Instructions</u>: Think of yourself as a student of Spanish and a language learner. Read each statement carefully. Consider how much you agree with each statement. There are no correct answers; only answers that best represent you.
**Note:* if you are a native Spanish speaker, this section does not apply to you.

[Motivated learning behavior]

51. If a Spanish course was offered at university or somewhere else in the future, I would like to take it.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

52. I think I am doing my best to learn Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

53. I am prepared to expend a lot of effort in learning Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

54. Compared to my classmates, I think I study Spanish relatively hard.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

55. I am working hard at learning Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

[Ideal L2 Self]

56. In the future, I can imagine myself as a person who uses Spanish in his/her daily life.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

57. In the future, I can imagine myself as a person who understands Spanish movies or music without subtitles.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

58. In the future, I can imagine myself as a person who has the ability to express his or her opinions or thoughts accurately in Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

59. In the future, I can imagine myself as a person who does not hesitate to speak Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

60. In the future, I can imagine myself as a person whose strength is being competent in Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

[Ought-to L2 Self]

61. It will have a negative impact on my life if I don't study Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

62. Learning Spanish is necessary because people surrounding me expect me to do so.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

63. I study Spanish because close friends of mine think it is important.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

64. I have to study Spanish, because, if I do not study it, I think my parents will be disappointed with me.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

65. My parents believe that I must study Spanish to be an educated person.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

[L2 learning experience]

66. I find Spanish really interesting.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

67. I would like to take more Spanish classes at university.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

68. I really enjoy learning Spanish.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

69. I always look forward to Spanish class.

(1)	(2)	(3)	(4)	(5)
totally disagree	somewhat disagree	neutral	somewhat agree	totally agree

70. Time seems to go by faster when I'm studying Spanish.

(1) totally disagree	(2) somewhat disagree	(3) neutral	(4) somewhat agree	(5) totally agree

Thank you for your responses to the Language Learner Profile Questionnaire! You may now proceed to your scheduled Zoom session with the researcher for the experimental portion of the study. If you have any questions or concerns at any time, please contact the researcher.
Thank you again and we look forward to meeting with you soon!

Appendix B: Spanish Proficiency Test (Task 1)

Part 1: Multiple Choice

Instructions: (presented in PPT slides)

You will see a series of sentences in Spanish, one per slide. Each of the following sentences contains a blank space (___) indicating that a word or phrase has been omitted. Read each sentence out loud carefully while selecting <u>one</u> of the four words provided to complete the sentence grammatically. Focus on both the meaning and structure of each sentence. If you are unsure, just guess and move on. There are 20 sentences in total.

1. Al oír del accidente de su buen amigo, Paco se puso ______.
   - alegre
   - fatigado
   - hambriento
   - desconsolado

2. No puedo comprarlo porque me ______ dinero.
   - falta
   - dan
   - presta
   - regalan

3. Aquí está tu café, Juanito. No te quemes, que está muy ______.
   - dulce
   - amargo
   - agrio
   - caliente

4. Al romper los anteojos, Juan se asustó porque no podía ______ sin ellos.
   - discurrir
   - oír
   - ver

- entender

5. Era una noche oscura sin ______.
   - estrellas
   - camas
   - lágrimas
   - nubes

6. ¡Qué ruido había con los gritos de los niños y el ______ de los perros!
   - olor
   - sueño
   - hambre
   - ladrar

7. ¡Cuidado con ese cuchillo o vas a ______ el dedo!
   - cortarte
   - torcerte
   - comerte
   - quemarte

8. Tuvo tanto miedo de caerse que se negó a ______ con nosotros.
   - almorzar
   - charlar
   - cantar
   - patinar

9. Compró ejemplares de todos los diarios, pero en vano. No halló ______.
   - los diez centavos
   - el periódico perdido
   - la noticia que deseaba
   - los ejemplos

10. Sus amigos pudieron haberlo salvado, pero lo dejaron ______.
    - ganar
    - parecer
    - perecer
    - acabar

11. Al lado de la Plaza de Armas había dos limosneros pidiendo ______.
    - pedazos
    - paz
    - monedas
    - escopetas

12. Siempre maltratado por los niños, el perro no podía acostumbrarse a ______ de sus nuevos amos.

- las caricias
- los engaños
- las locuras
- los golpes

13. ¿Dónde estará mi cartera? La dejé aquí mismo hace poco y parece que el necio de mi hermano ha vuelto a ______.

- dejármela
- deshacérmela
- escondérmela
- acabármela

14. En vez de dirigir el tráfico estabas charlando, así que tú mismo ______ del choque.

- sabes la gravedad
- eres testigo
- tuviste la culpa
- conociste a las víctimas

15. Posee esta tierra un clima tan propio para la agricultura como para _____.

- la construcción de trampas
- el fomento de motines
- el costo de vida
- la cría de reses

16. Aficionado leal de obras teatrales, Juan se entristeció al saber ______ del gran actor.

- del fallecimiento
- del éxito
- de la buena suerte
- de la alabanza

17. Se reunieron a menudo para efectuar un tratado, pero no pudieron ______.

- desavenirse
- echarlo a un lado
- rechazarlo
- llevarlo a cabo

18. Se negaron a embarcarse porque tenían miedo de______.

- los peces
- los naufragios
- los faros

- las playas

19. La mujer no aprobó el cambio de domicilio pues no le gustaba ______.
    - el callejeo
    - el puente
    - esa estación
    - aquel barrio

20. Era el único que tenía algo que comer, pero se negó a ______.
    - hojearlo
    - ponérselo
    - conservarlo
    - repartirlo

Part 2: Reading

Instructions: (presented in PPT slides)

In the following slides, you will read a short text about the life of a famous Spanish artist, Juan Miró. The text has been divided up over five slides and some of the words have been removed. Read out loud each segment of text *carefully* and choose <u>one</u> of the three words provided in parentheses that best completes the text. Focus on both the meaning and structure of each sentence. If you are unsure, just guess and move on. There are 12 blanks in total.

[Slide 1]

El sueño de Juan Miró

Hoy se inaugura en Palma de Mallorca la Fundación Pilar y Joan Miró, en el mismo lugar en donde el artista vivió sus últimos treinta y cinco años. El sueño de Joan Miró se ha ________________ (cumplido / completado / terminado)…

[Slide 2]

Los fondos donados a la ciudad por el pintor y su esposa en 1981 permitieron que el sueño se ___________ (inició / iniciara / iniciaba); más tarde, en 1986, el Ayuntamiento de Palma de Mallorca decidió ___________ (encargar / pedir / mandar) al arquitecto Rafael Moneo un edificio que ___________ (hubiera servido / haya servido / sirviera) a la vez como sede de la entidad y como museo moderno.

[*Slide 3*]

El proyecto ha tenido que ___________ (superar / enfrentarse / acabar) múltiples obstáculos de carácter administrativo. Miró, coincidiendo ___________ (por / en / con) los deseos de toda su familia, quiso que su obra no quedara expuesta en ampulosos panteones de arte o en __________ (voluntad / poder / favor) de coleccionistas acaudalados; por ello, en 1981, creó la fundación mallorquina…

[*Slide 4*]

Y cuando estaba __________ (al / en / a) punto de morir, donó terrenos y edificios, así como las obras de arte que en ellos __________ (habría / había / hubo).

[*Slide 5*]

El edificio que ha construido Rafael Moneo se enmarca en __________ (que / el que / lo que) se denomina "Territorio Miró", espacio en el que se han __________ (pretendido / tratado / intentado) de situar los distintos edificios que constituyen la herencia del pintor. El acceso a los mismos quedará __________ (disminuido / escaso / restringido) para evitar el deterioro de las obras.

Appendix C: stimuli for Speeded Oral Imitation Task (Task 2)

Instructions: (presented in PPT slides)

You will now listen to a series of statements in Spanish, one per slide. When you are ready, the researcher will play the short audio recording. You will listen to the recording and then indicate orally whether you agree or disagree with the content of the statement you just heard, based on your own personal preferences and opinions. Once you've responded whether you agree or disagree with the statement, you will repeat the statement you heard back to the researcher in *correct* Spanish to the best of your ability. *Please respond as quickly as possible with your first impression*. We will begin with a short practice. There are 24 sentences in total.

Practice:
1. [Me gusta correr todos los días.]
2. [Quiero vivir en un edificio en el centro.]

Linguistic variable distribution: all targets (24)
- 12 grammatical (6 Det-N focus; 6 N-Adj. focus);
- 12 ungrammatical (6 Det-N errors; 6 N-Adj. errors)

Grammatical (12):
- Det-N focus (6):
  - Masculine nouns (3):
    - overtly marked (2):
      1. El dinero es verde.
      2. Prefiero un espejo grande.

- non-overtly marked (1):
    - 3. Quiero hacer un viaje internacional.
- Feminine nouns (3):
  - overtly marked (2):
    - 4. Quiero ver una película interesante.
    - 5. Necesito una computadora más eficiente.
  - non-overtly marked (1):
    - 6. Prefiero tomar la leche regular.
- N-Adj. focus (6):
  - Masculine nouns (3):
    - overtly marked (1):
      - 7. Mi escritorio está bien organizado.
    - non-overtly marked (2):
      - 8. Cada parque bonito tiene árboles.
      - 9. Cada hotel es sucio.
  - Feminine nouns (3):
    - overtly marked (1):
      - 10. Mi casa es moderna.
    - non-overtly marked (2):
      - 11. Mi universidad es muy buena.
      - 12. Mi habitación está limpia.

Ungrammatical (12):

- Det-N error (6):
  - Masculine nouns (3):
    - overtly marked (1):
      - 13. Una museo elegante es mejor.
    - non-overtly marked (2):
      - 14. La café es saludable.
      - 15. Mañana tengo una examen importante.
  - Feminine nouns (3):
    - overtly marked (2):
      - 16. Prefiero ir a un playa canadiense.
      - 17. Prefiero comer un manzana verde.
    - non-overtly marked (1):
      - 18. El sal es verde.
- N-Adj. error (6):
  - Masculine nouns (3):
    - overtly marked (2):
      - 19. Cada queso es negra.
      - 20. Cada huevo es blanca.
    - non-overtly marked (1):
      - 21. Mi suéter es roja.

- Feminine nouns (3):
  - overtly marked (1):
    22. Cada montaña es bonito.
  - non-overtly marked (2):
    23. Mi nariz es bonito.
    24. Cada calle es largo.

Appendix D: stimuli for Speeded Auditory Grammaticality Judgement Task (AGJT) (Task 3)

Instructions: (presented in PPT slides)

You will listen to a series of short sentences in Spanish, one per slide. When you are ready, the researcher will play the short audio recording. You will listen to the recording and then indicate orally whether the sentence you heard is grammatically correct or incorrect. Please respond as quickly as possible with your first impression. We will begin with a short practice. There are 32 sentences in total.

Practice:
1. [Usted vas a correr en el parque.] *incorrect*
2. [Ella quiere vivir en el centro.] *correct*

Linguistic variable distribution:
targets (16) + distractors (16)
targets: 8 grammatical; 8 ungrammatical
- 8 grammatical (4 Det-N focus; 4 N-Adj. focus);
- 8 ungrammatical (4 Det-N errors; 4 N-Adj. errors)

Targets (16):
- Grammatical (8):
  - Det-N focus (4):
    - Masculine (2):
      - overtly marked (1):
        1. El barco es muy fuerte.
      - non-overtly marked (1):
        2. El reloj es grande.
    - Feminine (2):
      - overtly marked (1):
        3. La mesa es circular.
      - non-overtly marked (1):
        4. La luz está brillante.
  - N-Adj. focus (4):

- Masculine (2):
  - overtly marked (1):
    5. Mi plato está vacío.
  - non-overtly marked (1):
    6. Mi lápiz es amarillo.
- Feminine (2):
  - overtly marked (1):
    7. Mi silla está rota.
  - non-overtly marked (1):
    8. Mi comunidad es pequeña.

- Ungrammatical (8):
  - Det-N error (4):
    - Masculine (2):
      - overtly marked (1):
        9. La sombrero es verde.
      - non-overtly marked (1):
        10. La tren es eficiente.
    - Feminine (2):
      - overtly marked (1):
        11. El corbata es terrible.
      - non-overtly marked (1):
        12. El flor es elegante.
  - N-Adj. error (4):
    - Masculine (2):
      - overtly marked (1):
        13. Su baño está sucia.
      - non-overtly marked (1):
        14. Cada cine es sucia.
    - Feminine (2):
      - overtly marked (1):
        15. Mi escuela es bonito.
      - non-overtly marked (1):
        16. Mi clase de español es pequeño.

Distractors (16): (verb phases with subj-verb agreement & violations)

- Grammatical (8):
  17. Yo voy a la biblioteca.
  18. Tú eres un estudiante.
  19. Nosotros estamos estudiando.
  20. Carla toma café cada mañana.
  21. Ellos estudian español.
  22. Yo escribo una carta.
  23. Jorge come papas fritas.
  24. Yo bebo mucha agua.
- Ungrammatical (8):
  25. Tú va a la farmacia.

26. Nosotras son buenas amigas.
27. Pedro estás en la cafetería.
28. Yo toma muchas fotos.
29. Ella escriben cada día.
30. Ustedes comemos mucho.
31. Tú toman refresco en el cine.
32. Ana y Lola estudiamos juntas.

Appendix E: stimuli for Speeded Written Grammaticality Judgement Task (WGJT) (Task 4)

Instructions: (presented in PPT slides)

You will read a series of short sentences in Spanish, one per slide. When you are ready, the researcher will present the sentence. You will read the sentence and then indicate orally whether the sentence you read is grammatically correct or incorrect. Please respond as quickly as possible with your first impression. We will begin with a short practice. There are 32 sentences in total.

Practice:
1. Nosotros comen mucho arroz. *incorrect*
2. Él está en la biblioteca. *correct*

Linguistic variable distribution:
targets (16) + distractors (16)
targets: 8 grammatical; 8 ungrammatical
- 8 grammatical (4 Det-N focus; 4 N-Adj. focus);
- 8 ungrammatical (4 Det-N errors; 4 N-Adj. errors)

Targets (16):
- Grammatical (8):
    - Det-N focus (4):
        - Masculine (2):
            - overtly marked (1):
                1. Veo un cuaderno azul.
            - non-overtly marked (1):
                2. El avión es impresionante.
        - Feminine (2):
            - overtly marked (1):
                3. La cama es grande.
            - non-overtly marked (1):
                4. La diversion es importante.
    - N-Adj. focus (4):
        - Masculine (2):

- overtly marked (1):
  - 5. Tu laboratorio es moderno.
- non-overtly marked (1):
  - 6. Su menú es muy variado.

- Feminine (2):
  - overtly marked (1):
    - 7. Su pizarra es negra.
  - non-overtly marked (1):
    - 8. Tu conversación está aburrida.

- Ungrammatical (8):
  - Det-N error (4):
    - Masculine (2):
      - overtly marked (1):
        - 9. La vestido es elegante.
      - non-overtly marked (1):
        - 10. La autobús está tarde.
    - Feminine (2):
      - overtly marked (1):
        - 11. El prueba es fácil.
      - non-overtly marked (1):
        - 12. El lección es interesante.
  - N-Adj. error (4):
    - Masculine (2):
      - overtly marked (1):
        - 13. Su vuelo es muy larga.
      - non-overtly marked (1):
        - 14. Su jabón es blanca.
    - Feminine (2):
      - overtly marked (1):
        - 15. Tu ventana está abierto.
      - non-overtly marked (1):
        - 16. Mi llave es nuevo.

Distractors (16): (verb phases with subj-verb agreement & violations)

- Grammatical (8):
  - 17. Tú bebes mucha agua.
  - 18. Yo como al mediodía.
  - 19. Ella está escribiendo una novela.
  - 20. Jorge toma fotos.
  - 21. Nosotros estamos en el parque.
  - 22. Él es estudiante de español.
  - 23. Carla estudia inglés.
  - 24. Los niños van a casa.
- Ungrammatical (8):
  - 25. Nosotros están felices.

26. Ustedes estudiamos mucho.
27. Ella comes el pan.
28. Ella necesitas estudiar.
29. Tú debe dormir ahora.
30. Yo está aquí contigo.
31. Ellos va a la piscina.
32. Nosotras escriben la tarea.

Appendix F: stimuli for Self-paced Written GJT (untimed WGJT) (Task 5)

*Task stimuli are taken from Task 3 (8 tokens) & Task 4 (8 tokens)
Instructions: (presented in PPT slides)

You will read a series of short sentences in Spanish, one per slide. When you are ready, the researcher will present the sentence. You will read the sentence in Spanish and then respond orally in three different ways, as detailed below:

1. grammatical judgement of sentence: ✓GRAMATICAL ✕ NOT GRAMATICAL

2. How certain are you of your judgement?
0% 20% 40% 60% 80% 100%

3. How did you know that your response was correct?
A. I'm not sure; I just guessed.
B. It just sounds right.
C. I remembered a rule.

Please take your time and focus on the grammatical structure of the sentence you read. We will begin with a short practice. There are 16 sentences in total.

Practice:
1. Corro todos las días. *incorrect*
2. Vivo en un edificio alto. *correct*

Linguistic variable distribution: all targets (16)
8 grammatical; 8 ungrammatical
- 8 grammatical (4 Det-N focus; 4 N-Adj. focus);
- 8 ungrammatical (4 Det-N errors; 4 N-Adj. errors)

Targets (16):
- Grammatical (8):

- Det-N focus (4):
    - Masculine (2):
        - overtly marked (1):
            1. Veo un número grande.
        - non-overtly marked (1):
            2. El papel es importante.
    - Feminine (2):
        - overtly marked (1):
            3. La maleta es marrón.
        - non-overtly marked (1):
            4. La nacionalidad es esencial.
- N-Adj. focus (4):
    - Masculine (2):
        - overtly marked (1):
            5. Su vídeo es largo.
        - non-overtly marked (1):
            6. Mi semestre es largo.
    - Feminine (2):
        - overtly marked (1):
            7. Su calculadora es negra.
        - non-overtly marked (1):
            8. Su reservación está lista.

- Ungrammatical (8):
    - Det-N error (4):
        - Masculine (2):
            - overtly marked (1):
                9. La estadio es grande.
            - non-overtly marked (1):
                10. La béisbol es interesante.
        - Feminine (2):
            - overtly marked (1):
                11. El piscina es grande.
            - non-overtly marked (1):
                12. El nube está cerca.
    - N-Adj. error (4):
        - Masculine (2):
            - overtly marked (1):
                13. Su zapato está sucia.
            - non-overtly marked (1):
                14. Su maíz es amarilla.
        - Feminine (2):
            - overtly marked (1):
                15. Tu sandalia es bonito.
            - non-overtly marked (1):
                16. Mi carne está rojo.

Appendix G: Metalinguistic Awareness Exit Survey (Task 6)

Slide 1 -
Based on the tasks you've just completed, what grammatical structure do you think we are researching in this study?

Slide 2 –
Below are four sentences that are grammatically incorrect in Spanish.

1. Prefiero ir a un playa canadiense. [Det.-N error: feminine]
2. La tren es eficiente. [Det.-N error: masculine]
3. Mi nariz es bonito. [N-Adj. error: feminine]
4. Su baño está sucia. [N-Adj. error: masculine]

Questions:
Why do you think these sentences are grammatically incorrect?
What rule applies here? Can you explain it in your own words?

Appendix H: Task 2—Speeded Oral Imitation task (auditory) data summarized

	SPEAKER STATUS	Mean	SD	Variance
TASK 2 SCORE	Native Spanish	0.990	0.024	0.001
	Spanish Learner	0.673	0.217	0.047
TASK 2 COMPLETION TIME	Native Spanish	4.095	1.020	1.041
	Spanish Learner	5.047	1.170	1.368
TASK 2 TPR	Native Spanish	1.000	0.000	0.000
	Spanish Learner	0.932	0.100	0.010
TASK 2 COHERENCE SCORE	Native Spanish	1.000	0.000	0.000
	Spanish Learner	0.825	0.227	0.051
TASK 2 VAR.	Native Spanish	0.010	0.024	0.001

	Spanish Learner	0.181	0.087	0.007
TASK 2 SD	Native Spanish	0.038	0.091	0.008
	Spanish Learner	0.395	0.155	0.024
TASK 2 GRAMM SCORE	Native Spanish	1.000	0.000	0.000
	Spanish Learner	0.811	0.197	0.039
TASK 2 UNGRAMM SCORE	Native Spanish	0.980	0.050	0.003
	Spanish Learner	0.534	0.297	0.088
TASK 2 MASC. SCORE	Native Spanish	0.990	0.037	0.001
	Spanish Learner	0.697	0.253	0.064
TASK 2 FEM. SCORE	Native Spanish	0.990	0.037	0.001
	Spanish Learner	0.647	0.216	0.047
TASK 2 OVERT SCORE	Native Spanish	0.997	0.016	0.000
	Spanish Learner	0.744	0.236	0.056
TASK 2 NON-OVERT SCORE	Native Spanish	0.984	0.042	0.002
	Spanish Learner	0.599	0.233	0.054
TASK 2 DET-N SCORE	Native Spanish	0.997	0.016	0.000
	Spanish Learner	0.661	0.245	0.060
TASK 2 N-ADJ SCORE	Native Spanish	0.984	0.042	0.002
	Spanish Learner	0.683	0.214	0.046

Appendix I: Task 3—Speeded Auditory Grammaticality Judgment task (GJT) data summarized.

	SPEAKER STATUS	Mean	SD	Variance
TASK 3 SCORE	Native Spanish	0.954	0.056	0.003
	Spanish Learner	0.736	0.147	0.022
TASK 3 COMPLETION TIME	Native Spanish	2.948	0.693	0.480
	Spanish Learner	3.613	0.695	0.482
TASK 3 VAR	Native Spanish	0.043	0.047	0.002
	Spanish Learner	0.179	0.071	0.005
TASK 3 SD	Native Spanish	0.162	0.138	0.019
	Spanish Learner	0.410	0.107	0.012
TASK 3 TARGET SCORE	Native Spanish	0.963	0.094	0.009
	Spanish Learner	0.646	0.197	0.039
TASK 3 DISTR. SCORE	Native Spanish	0.885	0.057	0.003
	Spanish Learner	0.776	0.128	0.016
TASK 3 GRAMM SCORE	Native Spanish	0.966	0.083	0.007
	Spanish Learner	0.787	0.183	0.034
TASK 3 UNGRAMM SCORE	Native Spanish	0.960	0.152	0.023
	Spanish Learner	0.507	0.280	0.078
TASK 3 MASC. SCORE	Native Spanish	0.960	0.112	0.013
	Spanish Learner	0.691	0.214	0.046
TASK 3 FEM. SCORE	Native Spanish	0.859	0.075	0.006
	Spanish Learner	0.533	0.216	0.047
TASK 3 OVERT SCORE	Native Spanish	0.966	0.091	0.008
	Spanish Learner	0.637	0.246	0.060
TASK 3 NON-OVERT SCORE	Native Spanish	0.961	0.106	0.011

	Spanish Learner	0.656	0.194	0.038
TASK 3 DET-N SCORE	Native Spanish	0.832	0.129	0.017
	Spanish Learner	0.557	0.187	0.035
TASK 3 N-ADJ SCORE	Native Spanish	0.990	0.050	0.003
	Spanish Learner	0.665	0.237	0.056

Appendix J: Task 4—Speeded Written Grammaticality Judgment task (GJT) data summarized.

	SPEAKER STATUS	Mean	SD	Variance
TASK 4 SCORE	Native Spanish	0.982	0.027	0.001
	Spanish Learner	0.827	0.110	0.012
TASK 4 COMPLETION TIME	Native Spanish	2.655	0.571	0.326
	Spanish Learner	3.432	0.751	0.564
TASK 4 VAR	Native Spanish	0.018	0.026	0.001
	Spanish Learner	0.136	0.074	0.005
TASK 4 SD	Native Spanish	0.088	0.107	0.011
	Spanish Learner	0.343	0.139	0.019
TASK 4 TARGET SCORE	Native Spanish	0.968	0.054	0.003
	Spanish Learner	0.723	0.182	0.033
TASK 4 DISTR SCORE	Native Spanish	0.995	0.017	0.000
	Spanish Learner	0.932	0.085	0.007
TASK 4 GRAMM SCORE	Native Spanish	0.966	0.055	0.003
	Spanish Learner	0.894	0.116	0.014

TASK 4 UNGRAMM SCORE	Native Spanish	0.970	0.104	0.011
	Spanish Learner	0.554	0.322	0.104
TASK 4 MASC. SCORE	Native Spanish	0.877	0.048	0.002
	Spanish Learner	0.685	0.185	0.034
TASK 4 FEM. SCORE	Native Spanish	0.952	0.071	0.005
	Spanish Learner	0.676	0.208	0.043
TASK 4 OVERT SCORE	Native Spanish	0.990	0.050	0.003
	Spanish Learner	0.741	0.189	0.036
TASK 4 NON-OVERT SCORE	Native Spanish	0.947	0.071	0.005
	Spanish Learner	0.707	0.207	0.043
TASK 4 DET-N SCORE	Native Spanish	0.995	0.024	0.001
	Spanish Learner	0.749	0.210	0.044
TASK 4 N-ADJ SCORE	Native Spanish	0.942	0.108	0.012
	Spanish Learner	0.699	0.218	0.047

Appendix K: Task 5—Self-paced Written Grammaticality Judgment task (GJT) data summarized.

	SPEAKER STATUS	Mean	SD	Variance
TASK 5 SCORE	Native Spanish	0.990	0.022	0.001
	Spanish Learner	0.773	0.158	0.025
TASK 5 COMPLETION TIME	Native Spanish	3.518	1.487	2.210
	Spanish Learner	4.869	1.651	2.727
TASK 5 CERT. RATING	Native Spanish	0.984	0.041	0.002

	Spanish Learner	0.790	0.140	0.020
TASK 5 AVG KNW. SOURCE ATTR.	Native Spanish	1.204	0.238	0.056
	Spanish Learner	1.405	0.284	0.081
TASK 5 VAR	Native Spanish	0.010	0.022	0.001
	Spanish Learner	0.162	0.085	0.007
TASK 5 SD	Native Spanish	0.040	0.094	0.009
	Spanish Learner	0.374	0.150	0.023
TASK 5 GRAMM SCORE	Native Spanish	0.986	0.040	0.002
	Spanish Learner	0.884	0.142	0.020
TASK 5 UNGRAMM SCORE	Native Spanish	0.995	0.024	0.001
	Spanish Learner	0.662	0.235	0.055
TASK 5 MASC. SCORE	Native Spanish	0.995	0.024	0.001
	Spanish Learner	0.789	0.203	0.041
TASK 5 FEM. SCORE	Native Spanish	0.986	0.040	0.002
	Spanish Learner	0.757	0.161	0.026
TASK 5 OVERT SCORE	Native Spanish	0.990	0.033	0.001
	Spanish Learner	0.838	0.182	0.033
TASK 5 NON-OVERT SCORE	Native Spanish	0.990	0.033	0.001
	Spanish Learner	0.708	0.188	0.035
TASK 5 DET-N SCORE	Native Spanish	0.990	0.033	0.001
	Spanish Learner	0.794	0.167	0.028
TASK 5 N-ADJ SCORE	Native Spanish	0.990	0.033	0.001
	Spanish Learner	0.753	0.182	0.033

Appendix L: Task 6—Metalinguistic Awareness Exit Survey task data summarized.

	SPEAKER STATUS	Mean	SD	Variance
TASK 6: OVERT AWARENESS	Native Spanish	0.600	0.500	0.250
	Spanish Learner	0.678	0.470	0.221
TASK 6: COND. AWARENESS	Native Spanish	1.000	0.000	0.000
	Spanish Learner	0.944	0.230	0.053
TASK 6: TOTAL AWARENESS	Native Spanish	0.800	0.250	0.063
	Spanish Learner	0.811	0.296	0.088

Curriculum Vitae

Name:	Martha R. Black
Post-secondary Education and Degrees:	Saint John Fisher College Rochester, NY, USA 2006-2010 B.A.
	The University of Delaware Newark, DE, USA 2013-2015 M.A.
	Western University London, Ontario, Canada 2018-2023 Ph.D.
Honours and Awards:	Graduate Thesis Research Award Faculty of Arts and Humanities Western University 2021-2023
	Faculty of Arts and Humanities Dean's Entrance Scholarship Western University 2018
	National Geographic Young Explorer's Grant 2010
Related Work Experience	Instructor of English and Linguistics Aix Marseille Université (AMU), France 2021-present
	Instructor of Spanish Western University, Canada 2018-present
	Instructor of Spanish University of Alaska Anchorage, USA 2015-2018
	Instructor of Spanish University of Delaware, USA 2013-2015

Instructor of English and Academic Writing
Pontificia Universidad Católica Madre y Maestra, Dominican Republic
2010-2012

Selected Publications:

Black, M., Rato, A., & Rafat, Y. (accepted). Effect of perceptual training without feedback on bilingual speech perception: Evidence from approximant-stop discrimination in L1 Spanish and L1 English late bilinguals. *Journal of Monolingual and Bilingual Speech (JMBS)*. https://journal.equinoxpub.com/JMBS/index

Black, M. (2022). Multilingual advantage for adult acquisition of morphosyntax. Western Papers in Linguistics, 5(1). ISSN 2564-1786. https://ojs.lib.uwo.ca/index.php/wpl_clw/article/view/14566

Black, M. (2021). Multilingual advantage for adult learning of morphosyntax and the effect of processing modality. *WEFLA 2021 Conference Proceedings*. ISBN 978-959-7237-99-0. https://eventos.uho.edu.cu/index.php/ccm/cci10/paper/view/4702

Black, M. (2021). Variation and individual differences in how adults learn Spanish. *Inspiring Minds – Showcasing Western's Graduate Research, Scholarship and Creative Activity*. 141. https://ir.lib.uwo.ca/inspiringminds/141

Black, M. & O. Tararova. (2020). Adult acquisition of grammatical gender in instructed L2 Spanish and the role of metacognition. *Proceedings of the 2020 annual conference of the Canadian Linguistic Association.* London, Canada. Online. https://cla-acl.artsci.utoronto.ca/wp-content/uploads/actes-2020/Black_Tararova_CLA-ACL2020.pdf

Black, M. (2020). Perception of similar sounds as a function of language dominance in late bilinguals of English and Spanish. In E. Rosales (Ed.), *Proceedings of the 8th Meeting on Language Teaching* (MeLT) (pp. 20-32). Montréal: UQÀM. https://sites.grenadine.uqam.ca/sites/melt/fr/edition2020/documents/get_document/30

Black, M., Joanisse, M.F., & Rafat, Y. (2020). Language dominance modulates the perception of Spanish approximants in late bilinguals. *Languages*, 5(1), 7. https://doi.org/10.3390/languages5010007

Black, M. (2019). Una aplicación pedagógica del Perfil del Lenguaje Bilingüe para promover la autoconciencia lingüística en el aprendizaje de L2. *WEFLA 2019 Conference Proceedings*. ISBN: 978-959-7237-34-1https://wefla2019.sciencesconf.org/data/37.pdf

Slemp, K., **Black, M.**, Cortiana, G. (2020). Reactions to gender-inclusive language in Spanish on Twitter and YouTube. *Proceedings of the 2020 annual conference of the Canadian Linguistic Association*. London, Canada. Online. https://cla-acl.artsci.utoronto.ca/wp-content/uploads/actes-2020/Slemp_Black_Cortiana_CLA-ACL2020.pdf

Tararova, O., **M. Black**, Q. Wang, K. Blong. (2023). Gender Agreement in L3 Spanish Production among Speakers of Typologically Different Languages. *Languages* 8(18). https://doi.org/10.3390/languages8010018