8-23-2024 3:00 PM

# Impact of Passive Second Language Exposure on Word Segmentation and Word Mapping

Amiya S. Aggarwal, *Western University*

Supervisor: Batterink, Laura, *The University of Western Ontario*
A thesis submitted in partial fulfillment of the requirements for the Master of Science degree in Psychology

Abstract

Research in statistical learning using artificial languages has suggested that passive exposure to linguistic patterns can guide word segmentation and word mapping. It is unknown whether this type of unsupervised learning can scale up to support these aspects of learning in a natural language. Our study exposed monolingual English speakers to either English or Italian podcasts for an hour daily for 21 days, collecting behavioural and EEG data during both the pre- and post-exposure period. Our behavioural measure tested if L2 exposure would lead to improved word mapping for words with high phonotactic probabilities. Our EEG measure tested whether the L2 exposure would improve word segmentation, as indexed by the TRF to word onsets. Neither of these measures detected significant changes in L2 processing as a function of passive exposure. Accounting for additional baseline acoustic features in the future may clarify the effects of L2 exposure on moment-by-moment L2 processing.

**Keywords**:

Statistical learning, language acquisition, second language learning, word mapping, word segmentation, temporal response functions, encoding models, EEG

# Summary for Lay Audience

It is common for new language learners to take second language classes, learn a large number of vocabulary items, and still be unable to pick out the words in fluent, seemingly rapid-fire speech produced by a native speaker. However, as many language learning applications encourage, listening to a language in the form of podcasts or audiobooks can boost comprehension. The goal of this research was to test if passively listening to a second language is sufficient to help monolingual English speakers find word boundaries and whether familiarity with the sound patterns of the second language makes it easier to treat common words from that language as labels for new objects. We asked participants to listen to either English or Italian podcasts for an hour a day and compared their performance on two distinct measures, a Word Mapping task and a Continuous Listening task. The Word Mapping task compared how well participants could make word-object associations for three word types – common words (typical sound patterns of Italian), rare words (atypical sound patterns for Italian) and non-words (atypical sound patterns for Italian). The Continuous Listening task recorded participants' neural response to word onsets in Italian speech. In contrast to our expectations, we did not find a measurable impact of the three weeks of exposure to a second language on either the Word Mapping or Continuous Listening task. However, on the word mapping measure, we found some evidence that all participants became more sensitive to sound patterns in the second language from session 1 to session 2. On the Continuous Listening task, it is possible that additional analyses which model other features of speech may reveal more subtle changes occurring during online speech processing. In the future, this line of research could uncover new ways to facilitate second language learning in adults by exploring methods other than traditional classroom teaching.

# Co-authorship Statement

I, Amiya Aggarwal, wrote this thesis independently.

My supervisor, Dr. Laura Batterink contributed to developing the research question, study design, and revisions to the thesis.

Rae Hoeppner contributed to selecting, trimming, editing podcasts and embedding words into the podcast audio files. They calculated the phonotactic scores for all selected test items used in the word mapping task, created orthographic representations for the non-words and programmed the continuous listening experiment.

Dr. Aaron Gibbings contributed significantly to all aspects of EEG data analysis. Specifically, modifying and creating custom functions to convert raw EEG files into CND format, stimulus alignment, creating figures and confirming analyses.

# Acknowledgements

I would like to thank my supervisor, Dr. Laura Batterink for being an incredibly supportive mentor throughout this entire process. Her positive energy and enthusiasm for the project motivated me to push myself and explore new ideas. I would also like to thank my advisors, Dr. Debra Jared and Dr. Yasaman Rafat for their encouragement and guidance.

The work presented here would not have been possible without the contributions of Rae Hoeppner, who has been a friend, colleague and support throughout this project, and Dr. Aaron Gibbings, who patiently guided me through working with EEG and MATLAB.

I would also like to thank the rest of the CNLLlab, particularly my undergraduate assistants who helped with hundreds of hours of data collection: Chelsea Ingham, Maream Al-Rajab, and Shveta Suresh.

And finally, I would like to express my gratitude to my family and friends for their constant support.

# Table of Contents

# List of Figures

# List of Appendices

# Chapter 1

## 1 Introduction

Studies of language processing have shown that the neural response to continuous speech depends on linguistic proficiency (Ihara et al., 2021). Fluent listeners of a language can effortlessly extract meaning from speech, whereas listeners who are less familiar with a language may not glean any semantic content from the auditory input. However, they may still be able to encode some of the acoustic features, such as the speech envelope or changes in sound intensity (Brodbeck et al., 2024; Gillis et al., 2021; Zou et al., 2019).

Familiarization with a new language is accompanied by a shift in auditory processing, such that surface level acoustic features are thought to become less salient with increased proficiency (Brodbeck et al., 2024; Di Liberto et al., 2021; Pérez-Navarro et al., 2024). In past studies that have related learners' neural response to their linguistic proficiency, the second language (L2) learners typically had at least some formal classroom training (Di Liberto et al., 2021; Ihara et al., 2021). However, it is not known whether mere exposure to a spoken L2 is sufficient to induce neural changes during language processing in adult learners, or whether neural changes can only result from more intentional, explicit forms of learning and instruction. Language proficiency also affects whether listeners are able to adapt word segmentation strategies while processing continuous L2 speech (Gilbert et al., 2021). Therefore, unfamiliar listeners may not be able to consistently detect word boundaries, which is a necessary component for the neural response to word onsets (Karunathilake et al., 2023).

At the behavioural level, language exposure enables listeners to develop a sensitivity to regularities in sound patterns (phonotactics), eventually supporting the ability to learn word-forms and grammar (Kittleson et al., 2010; Plante & Gómez, 2018; Romberg & Saffran, 2010). This familiarity with sound patterns and prototypical word structures can further influence the ability to associate sound sequences to objects, known as word mapping (Graf Estes et al., 2007; Hay et al., 2011; Mirman, Magnuson, et al., 2008). This thesis will investigate the extent to which passive L2 exposure facilitates word segmentation in natural L2 speech, and whether generalized word-form knowledge can be acquired and further promote the learning of object-label associations.

## 1.1 Word Segmentation Through Statistical Learning

A critical first step to language acquisition is *word segmentation* – the ability to discover individual words in continuous speech (Saffran, Newport, et al., 1996). However, the process of word segmentation can be challenging due to the lack of consistent acoustic cues that indicate word boundaries. Speech is composed of an overlapping combination of language-specific cues to word boundaries, such as lexical stress patterns (indicated through higher pitch, longer duration, and increased volume of stressed syllables), and language-general cues, such as syllable co-occurrence patterns (i.e. transitional probabilities) and utterance boundaries (Sahni et al., 2010). Becoming sensitive to the probabilistic regularities of speech across repeated instances may enable the reliable extraction of word boundaries through the process of *statistical learning* (Saffran, Aslin, et al., 1996). Statistical learning refers to the process of extracting patterns from the environment simply through exposure to input, without explicit training, effort, or intention (Plante & Gómez, 2018; Romberg & Saffran, 2010).

Word segmentation through statistical learning is a phenomenon that has been studied extensively using miniature artificial languages (e.g., Isbilen & Christiansen, 2022; Romberg & Saffran, 2010; Saffran et al., 1997; Saffran, Newport, et al., 1996). Artificial language studies have focused on the universal segmentation cue of *transitional probabilities,* which are calculated as the likelihood of a specific syllable occurring immediately after the preceding one (Saffran, Newport, et al., 1996). Statistical learning studies using artificial languages typically present a continuous stream of syllables made up of 4-6 randomly concatenated syllable triplets, or “words” (e.g. *bidakupadotigolabubidaku…*). They are designed such that syllables within a word are always found together (TP = 1), but those between words are less likely to be found together (e.g. TP = .33), though this varies depending on the number of words (Isbilen & Christiansen, 2022).

Research suggests that infants are able to leverage statistical learning mechanisms to become sensitive to language-general cues which then allow them to discover more language-specific cues in their native language (Sahni et al., 2010). Through a series of seminal studies, Saffran and colleagues (Saffran, Aslin, et al., 1996; Saffran et al., 1997; Saffran, Newport, et al., 1996) demonstrated that adults, children and infants are able to pick up on regularities in structured input to identify words in a continuous syllable stream made up of trisyllabic words, with no pauses or other acoustic cues indicating word boundaries. After only 2 minutes of exposure to this artificial language, infants were able to differentiate previously heard sequences from new ones. This discrimination ability was demonstrated through a novelty preference in a looking-time procedure in which the infants spent significantly longer listening to non-words than those that had occurred previously in the artificial speech stream (Saffran, Aslin, et al., 1996). Whereas infant studies of statistical learning leverage looking-time procedures, studies with adults and children typically use explicit measures to assess learning, such as asking participants to rate the familiarity of test items on a Likert-type scale (e.g., Batterink & Paller, 2017, 2019; Liu et al., 2023), or to choose between words and non-words (recombined syllables from the language) on a two-alternative forced choice (2AFC) recognition task (e.g., Batterink et al., 2015; Elmer et al., 2021; Saffran et al., 1997; Saffran, Newport, et al., 1996). Based solely on the transitional probabilities between neighbouring syllables, participants in these studies were able to rate the syllable sequences making up 'true' words as more familiar than non-word foil items, and also correctly identify them on the 2AFC task.

While the previously discussed studies have used explicit measures, statistical learning in adults can also be measured through more implicit measures. Researchers have found evidence for statistical learning using a target detection task, in which participants are required to make keypress responses to target syllables embedded in shortened excerpts of the artificial language stream. Participants in this task are thought to be able to predict later syllables in a word using the preceding syllables, and thus responding to later-occurring syllables (2nd and 3rd positions) faster than word-initial syllables (Batterink et al., 2015; Franco et al., 2015; Lukics & Lukács, 2021; Poulin-Charronnat et al., 2017). Faster reaction times to later syllables within a word were found even in participants who did not show evidence of learning as assessed on the explicit 2AFC recognition task (Batterink et al., 2015). Another more implicit measure of statistical learning is neural entrainment. EEG studies using artificial languages show that as participants are exposed to a syllable stream composed of trisyllabic words, their brain waves reflect a shift from processing individual syllables to segmenting trisyllabic units, as revealed by an increase in neural entrainment to the word frequency over time (Batterink & Paller, 2017; Buiatti et al., 2009; Elmer et al., 2021; Kabdebon et al., 2015; Sjuls et al., 2023).

Importantly, statistical learning is a powerful mechanism that occurs in response to linguistic input even without focused attention to the input or explicit effort to find patterns. An artificial language study found that both first grade children and adults successfully learned the words while hearing the speech stream in the background (Saffran et al., 1997). Participants in this study were asked to create computer illustrations while a 21-minute tape of the artificial language was being played in the room. Despite not being instructed to pay attention to the stream or being informed that the syllables they heard were part of a nonsense language, participants were able to recognize ‘true’ words as more like the tape they heard than non-word foils. This study led to the conclusion that passive exposure is sufficient for statistical learning of syllable sequences, and raises the possibility that incidental learning may play a role in natural language acquisition. Likewise, another study by Batterink and Paller (2019) provides further evidence that focused attention is not necessary for statistical learning. In the study, participants were exposed to an artificial language stream while engaged in a cognitively demanding (visual 3-back) task. Despite divided attention, they still showed evidence of word learning on various implicit and explicit measures: faster target detection for predictable syllables, higher familiarity ratings for words from the exposure stream and a gradual increase over time in neural entrainment to the word frequency. These results, which show that passive exposure leads to word segmentation in an artificial language, support the possibility that statistical learning mechanisms may benefit natural L2 acquisition from ambient exposure.

## 1.2 Word Segmentation Facilitates Word Mapping

The previously described studies of statistical learning (e.g., Saffran, Aslin, et al., 1996; Saffran et al., 1997; Saffran, Newport, et al., 1996) have shown that infants, children and adults are capable of segmenting words from continuous speech based on the distributional cues (for a review see: Isbilen & Christiansen, 2022). Building on these findings, Graf Estes and colleagues (2007) aimed to determine whether statistical word segmentation leads to a word-like representation that could potentially facilitate subsequent stages of language learning, such as vocabulary acquisition. To investigate this question, the authors used an object-label association task, or "word-mapping" task, developed by Stager and Werker (1997). They exposed 17-month-old English-learning infants to 2.5 minutes of a fluent speech stream composed of bisyllabic words, and then tested whether the segmented 'high probability sequences' were treated as candidate words. In a two-part experiment, words from the stream were compared to either non-words made up of unfamiliar syllables that were not part of the initial speech stream, or part-words containing syllables spanning word boundaries. Infants were habituated to two object-label pairings and then tested on how long they looked at trials with the same object-label pairing versus a switched one. Longer looking times for switch trials were used as an indicator of learning. Results showed that infants more readily mapped words to objects than non-words or part-words. The reduced ability of infants to link part-words to novel object labels showed that it is not only familiarity to individual syllable units, but the predictability of sound sequences that promotes subsequent word learning. Overall, these results indicate that statistical segmentation does provide a privileged representation of candidate words that can then facilitate subsequent stages of language learning.

A subsequent study aimed to extend these findings to adults, testing whether there is a direct link between statistical word segmentation and word learning in adult learners (Mirman, Magnuson, et al., 2008). In this study, adult participants were first exposed to an artificial language made up of bisyllabic words and then asked to complete an associative learning task. This task required participants to indicate which of two images corresponded to an auditorily presented word via a key press response. While participants were initially required to guess, over time they could make use of feedback to learn the correct object-label pairings. Although participants were able to learn all three word types, they were slower to learn the associations of objects to ‘part-words’ – those crossing word boundaries – than to words that occurred in the stream, and even to non-words made up of entirely different syllables. The findings of this study link statistical word segmentation to word learning in adults, and demonstrate that labels inconsistent with the statistics of the speech stream are harder to map to subsequent referents.

Building on the finding that infants map ‘words’ more readily than ‘part-words’ to objects (Graf Estes et al., 2007), researchers tested whether a similar phenomenon is observed using natural language stimuli (Hay et al., 2011; Shoaib et al., 2018). In these studies, English learning infants were familiarized with infant-directed speech in the form of naturally spoken Italian sentences. These sentences contained target words that were bisyllabic nouns and had carefully controlled transitional probabilities by ensuring that that syllables from high transitional probability words (TP = 1) never occurred elsewhere in the corpus, and that only the first and not the second syllable from low transitional probability words (TP = 0.33) occurred several times outside the target word. The authors found an advantage for word mapping of high transitional probability words compared to low transitional probability words, as shown by looking-time differences. This important finding suggests that, through either word segmentation or sound sequence familiarity, exposure to natural spoken language may facilitate further aspects of word learning.

## 1.3 Learning From Natural Linguistic Exposure

While the evidence supporting the role of statistical learning in speech segmentation is robust (Isbilen & Christiansen, 2022), nearly all of this evidence is in the context of miniaturized, artificial languages.  Unlike language input in the real world, the artificial languages used in these studies make use of an extremely limited stimulus set, contain frequent word repetitions, and intentionally contain little variability in word features (i.e. parts of speech, number of syllables, etc.) and prosody. This raises the question as to whether statistical learning can scale up to support aspects of natural language learning, or whether it is a phenomenon that is observed only in limited and artificial experimental contexts.

Some evidence in support of the scalability of statistical word segmentation comes from an artificial language learning study in which participants successfully learned words from unsegmented input consisting of 1000 novel words (Frank et al., 2013). This study required participants to passively listen to audio recordings of the artificial language for 1 hour daily over the course of 10 days. In addition to using a much larger language set than most artificial word segmentation studies (1000 words as opposed to only four or six), the authors also introduced additional variability in the language by varying word length, the frequency of words (following a Zipfian distribution), and the insertion of pauses as indicators of sentence boundaries, thereby mimicking some aspects of natural language. Participants in this study were not only able to use distributional cues to segment words from this more complicated language, but also were able to remember some of these words when tested three years later.

Further evidence that statistical learning mechanisms extend beyond the context of limited artificial language studies comes from a study of English-learning 8-month-olds (Pelucchi et al., 2009). The infants were familiarized with a series of Italian sentences containing bisyllabic target words. When tested later using a looking time paradigm, they were successfully able to discriminate familiar words, as shown by longer looking times, from novel words that did not occur in the sentences. This demonstrated that infants are capable of recognizing words from fluent L2 speech. In a second experiment, syllables from the familiar and novel words appeared an equal number of times in the familiarization sentences. Results indicated that infants showed a significant preference for familiar words over novel words, indicating that they were tracking syllable sequences rather than only individual syllables in natural speech. In a final experiment, the researchers compared two types of familiar words, those with high transitional probabilities and those with low transitional probabilities. Despite both word types occurring the same number of times, infants still showed a preference for the high transitional probability sequences over low transitional probability sequences, indicating that they tracked transitional probability information. Taken together, these results suggest that statistical learning can allow infants to learn the transitional probabilities of syllable sequences in a natural language.

There is also evidence that natural linguistic exposure alone can produce sensitivity to some aspects of an L2 in adult learners. In one study, English-speaking adults in New Zealand, who had a lifetime of ambient but non-intentional exposure to Māori, showed evidence of a proto-lexicon for Māori, as indexed by their ability to distinguish real Māori words from Māori-like non-words (Oh et al., 2020). In another test, designed to probe generalized phonotactic knowledge, participants were asked to provide well-formedness ratings for a series of non-words that differed in their 'word-like-ness'. The native Māori-speaking and English-speaking New Zealanders were equally accurate, and both significantly outperformed a group of American participants without previous exposure to the Māori language. Certain universal aspects of language may have allowed even the control group (American English speakers) to identify phonotactically illegal sound combinations, but these only accounted for minimal sensitivity to phonotactics. This study provides evidence that sensitivity to prototypical aspects of L2 word structure can be acquired through ambient exposure.

While the English-speaking New Zealanders had a lifetime of exposure to Māori, other studies have shown that participants can demonstrate some word-form knowledge in as little as 7 minutes of exposure to a foreign language. A study by Gullberg and colleagues (2010) found that exposing Dutch natives to 7-14 minutes of controlled, but naturalistic audiovisual input in the form of a Mandarin weather report was sufficient for participants to develop some phonotactic knowledge of Mandarin. The authors found that participants were better able to recognize familiar bisyllabic words that contain high transitional word-internal probabilities, as compared to monosyllabic words that lack any word-internal transitional probabilities. Word frequency also influenced learning, as participants showed better recognition of bisyllabic test items that occurred eight times than those that only occurred twice. In addition, participants were able to generalize phonotactic knowledge to new items, correctly rejecting monosyllabic pseudowords that contained syllable structure violations. While control participants performed at chance, participants exposed to Mandarin showed an increasing ability to reject the illegal syllables as exposure increased from 7 to 14 minutes. Most importantly, the test syllable type (CVC: e.g., gam) that was illegal in Mandarin is acceptable in Dutch, which shows how quickly L2 phonotactic knowledge can be acquired “from scratch”. Another study by Kittleson et al. (2010) also showed that adults can acquire knowledge of L2 phonotactics. In this study, non-Norwegian speaking participants who had recently moved to Norway listened to 7.2 minutes of Norwegian sentences containing bisyllabic test items and were then tested on their ability to discriminate Norwegian words from non-word foils. Across two experimental testing sessions, participants correctly rejected non-words, although their actual recognition of true words was at chance levels. Importantly, English-speaking control participants who hadn’t been exposed to Norwegian were not able to successfully reject non-words, ruling out contributions of baseline sensitivity to performance. This study provides further evidence that adults can develop a sensitivity to phonotactic regularities from quite brief exposure to fluent L2 speech.

In addition to acquiring phonotactic knowledge through targeted exposure to L2 input (Gullberg et al., 2010; Kittleson et al., 2010), naive listeners also appear to be able to extract phonotactic regularities from passively hearing an L2 (Alexander et al., 2023). In this study, English-speaking participants completed a familiarity rating task before and after a two-week exposure period during which they listened to Italian podcasts (L2 group) or English podcasts (controls) in the background for an hour a day, while going about everyday activities. From the first to the second testing sessions, participants in the L2 group significantly improved their ability to discriminate Italian words and non-words as compared with controls. This finding indicates that passive exposure to a second language leads to word form knowledge of an L2. In principle, this learning may be driven by an increase in sensitivity to various word features, including transitional probabilities, phonetic patterns, and word-stress patterns.

## 1.4 Modeling Neural Changes in Language Processing

Natural languages lack the carefully controlled transitional probabilities and word frequencies found in artificial languages, and thus natural language processing must be studied using different experimental approaches. Predictive modelling is a common approach that researchers have used to look at various aspects of natural language processing (Hamilton & Huth, 2020; Sassenhagen, 2019). Since behavioural measures of learning such as recall or familiarity judgement tasks often take place after learning has occurred, it is difficult to determine whether task performance reflects a change in the moment-by-moment processing of speech or on "offline" processes that rely on strategic processing, introspection or application of first language (L1) knowledge. Temporal response functions (TRFs) predict neural data from a set of time-aligned stimulus features (Crosse et al., 2016, 2021). By factoring in spatial components and averaging across many instances where they may overlap, it is possible to isolate the unique neural response elicited by specific stimulus features, such as phoneme onsets, or semantic dissimilarity. Compared to traditional ERP studies that require the use of discrete or isolated utterances, modeling TRFs to relatively long segments with continuous auditory input allows for a response to individual language features in a more ecologically valid context of uninterrupted speech (Crosse et al., 2021; Hamilton & Huth, 2020).

Neural tracking refers to the phenomenon in which brain responses synchronize with particular properties of sensory input and can be applied to understand speech processing (Gillis et al., 2021; Zou et al., 2019). By recording EEG, we can measure the neural response to natural speech and then align this neural response to acoustic, lexical, sub-lexical and other linguistic features in the input. Studies of neural tracking (Brodbeck et al., 2018; Di Liberto et al., 2021; Gillis et al., 2021) have shown that the neural response to speech can be reliably predicted based on various features, including acoustic features (e.g., speech envelope and acoustic onsets) and linguistic features (e.g., phoneme and word processing). Linguistic features, such as word- and phoneme-surprisal, have been shown to uniquely contribute to the recorded neural signal, beyond contributions of low-level acoustic processing (Gillis et al., 2021).

Using the mTRF approach, researchers have demonstrated that language proficiency affects the neural response to different linguistic features. In a study of native Japanese speakers with varying levels of L2 proficiency in English (Ihara et al., 2021), researchers could reliably decode the participants' proficiency levels based on their neural response to continuous English speech. As familiarity and expertise increased, more complex linguistic features, such as word class, became a better predictor of the neural response. In contrast, for beginners, or individuals with limited language ability, the basic acoustic features were the best predictors of the EEG response (Di Liberto et al., 2021; Pérez-Navarro et al., 2024). This finding indicates that the complexity of speech processing, as assessed through neural measures, varies as a function of language experience.

While some of the previously mentioned studies (Brodbeck et al., 2024; Gillis et al., 2021) incorporate word onsets into a model of baseline acoustic processing, word onsets have their own unique neural representation. In a study that looked at event related brain potentials (ERPs) in response to continuous words in an artificial language stream, researchers found a greater N100 response to the onset of segmented words after training, but only in participants who showed greater behavioural evidence of word learning (Sanders et al., 2002). These results show that the N100 response can index word segmentation even in the absence of acoustic cues. Supporting the idea that the N100 response can be used as a marker for word segmentation, an ERP study (Sanders & Neville, 2003a) found that word-initial syllables elicited a larger N100 response than word-medial syllables, even in sentences lacking semantics or syntactic structure. While the previous study was done in native speakers, the researchers also followed up with late L2 learners but found that they do not show this effect, indicating that their language processing differed from that of native speakers (Sanders & Neville, 2003b).

The neural response to word onsets has also been demonstrated in fully natural languages. Using the mTRF approach to model MEG data, Brodbeck and colleagues (2018) examined the neural response to word onsets as a measure of word segmentation in continuous L1 speech. A strong response to word onsets was observed with a peak at

103 ms showing a similar response to the N100 seen in ERP studies. This early response is a perceptual feature that is dependent on either top down or automatic linguistic processing and requires knowledge of word boundaries. If statistical learning mechanisms can facilitate word segmentation in a novel language, we would expect sensitivity to word onsets to be especially sensitive to L2 experience as this response is an index of linguistic but not acoustic processing (Karunathilake et al., 2023; Romberg & Saffran, 2010; Sanders & Neville, 2003). The neural response to word onsets would allow us to see if early L2 learners can segment the speech stream at the perceptual level.

In sum, previous research has shown that there is a neural response to the various levels of acoustic and language processing, and that these evolve as a function of linguistic proficiency (Di Liberto et al., 2021; Ihara et al., 2021). Even when exposed to a completely unfamiliar foreign language, individuals are able to track the acoustic features of the input (Brodbeck et al., 2024; Zou et al., 2019). In contrast, neural tracking of linguistic features such as word onsets depends on speech intelligibility (Karunathilake et al., 2023), and can be used as a measure of the moment-by-moment processing of linguistic input. In the current study, we will use the mTRF approach as a tool to investigate whether passive L2 exposure can lead to an increased sensitivity to word boundaries, as indexed by the neural response to word onsets.

## 1.5 The Present Study

The present study investigates the impact of exposure to a natural, spoken L2 on word segmentation, thereby testing whether statistical learning mechanisms extend to real-world language acquisition. We included two separate tasks – a Word Mapping task and a Continuous Listening task – in order to assess the development of phonotactic knowledge and sensitivity to word onsets through ambient exposure to spoken language. Participants completed the Word Mapping task and Continuous Listening task before and after a 3-week listening period. The listening period required them to listen to either Italian podcasts (L2 group) or English podcasts (Control group) in the background while going about routine activities for 1 hour every day. We hypothesized that even from limited passive exposure to language stimuli, listeners would extract regularities of the language that aid in word segmentation.

The Word Mapping task, previously used with artificial languages and limited natural languages (Mirman, Magnuson, et al., 2008), was adapted to our test language - Italian. Past studies of word mapping have shown that familiar words, once segmented, are mapped faster to their visual referents than non-words (Graf Estes et al., 2007; Hay et al., 2011; Mirman, Magnuson, et al., 2008; Shoaib et al., 2018). We hypothesized that familiarity with L2 specific sound patterns (phonotactics) would facilitate word mapping, such that the association between new prototypical L2 words to novel objects would be facilitated. Therefore, we predicted that, post exposure, participants in the L2 group would more readily map words with L2 consistent phonotactics to their associated objects compared to L2 inconsistent words. In contrast, we would not expect a difference between word types for participants in the control group.

The Continuous Listening task involved having participants listen to 40 minutes of Italian speech. The mTRF approach was used to test if passive L2 exposure facilitates speech segmentation, as assessed by participants' EEG response to word onsets. We hypothesized that listeners would not initially be sensitive to word boundaries in continuous L2 speech, but that a period of L2 exposure would increase their sensitivity to acoustic and phonotactic cues to word boundaries, as supported through statistical learning. Therefore, we predicted that at baseline all participants would only track acoustic information and not show a significant response to word onsets. Furthermore, only participants in the L2 group (exposed to Italian), would develop a sensitivity to word onsets post-exposure, as measured by an increase in the prediction correlations for encoding models including this feature.

# Chapter 2

# 2 Methods

## 2.1 Participants

Our final sample consisted of 45 self-reported monolingual English-speaking adults between the ages of 17-35 years ($M$ = 21.18, $SD$ = 2.59; 28 female). Participants were recruited through emails sent to staff and students associated with Western University and through flyers posted in the larger London community. A total of 59 participants were initially recruited and completed the first testing session. However, 14 participants were excluded for not complying with the podcast listening protocol ($n$ = 1), for inability to complete the second testing session ($n$ = 3), or for missing more than two consecutive daily podcast listening surveys ($n$ = 10); this latter group was dropped from further participation in the study at this point. We initially recruited 32 participants for the Italian listening (L2) group and 27 participants for the English listening (control) group. After exclusions, our final sample consisted of 27 participants in the L2 group and 18 in the control group. We intentionally oversampled the L2 group in order to prioritize examining the effect of L2 exposure on our experimental outcomes (i.e. tests involving the effect of session within this group).

In the final sample, participants all spoke English as their primary language and reported no significant experience with any other language, although some ($n$ = 6) reported taking basic French classes until 9$^{th}$ grade and others ($n$ = 2) also reported some limited experience with other languages (e.g., Arabic, Dutch and Portuguese).

## 2.2 Tasks and Stimuli

### 2.2.1 Listening Podcasts

A total of 21 Italian podcasts (assigned to the L2 group) and 21 English podcasts (assigned to the control group) were compiled for the listening portion of the study.

These consisted of segments from various podcast episodes concatenated together using the *Audacity(R)* software and edited to include embedded 'secret' words. Each file was approximately 1 hour long (mean length: 60.49 minutes, range: 53.6 - 66.5 minutes). The variability in segment length allowed for distinct story segments to conclude, creating a more natural narrative. Each audio file in both Italian and English listening conditions contained 3-5 'secret' everyday English words (e.g. "yellow", "bridge") that were used as attention and compliance checks. The secret English words were preceded by a chime so that participants would be able to easily distinguish them from the main podcast.

The Italian listening stimuli were taken from the 'Advanced' podcast series on the website *News in Slow Italian*, a resource for Italian language learners. These podcasts consisted of single speakers discussing the news and current events in a rehearsed but natural manner. The audio files that were used for the tasks were played at a natural speaking rate as confirmed by two native Italian speakers. The English listening podcasts were selected to match the Italian stimuli as closely as possible, and also consisted of single speakers speaking in a rehearsed but natural manner. The English podcasts contained a combination of clips from three podcast series: *Times the Brief, The Lazy Genius*, and *Newsworthy*. The *Newsworthy* and *Times the Brief* series presented news stories in a similar rehearsed but natural manner by a female and male speaker respectively. One difference in the content of the English podcasts was that *The Lazy Genius* podcast had slightly different subject material, and discussed ways to improve productivity rather than current events. A variety of podcasts were included to ensure that each hour-long file would have segments with both male and female voices.

### 2.2.2 Word Mapping Task

The Word Mapping task involved learning the associations between 12 words and 12 images. There were two counterbalanced experiment versions (A, B) to enable distinct stimulus sets for the two experimental sessions. Each version was further subdivided into 12 counterbalanced sub-versions. Across the counterbalanced sub-versions, the 12 target words were paired with 12 target images exhaustively to avoid any idiosyncratic word-image pair effects across participants.

### Visual Stimuli

Each experiment version (A/B) had 12 images that were randomly selected from a larger stimulus set. Following previous word mapping studies (Graf Estes et al., 2007; Hay et al., 2011; Mirman, Magnuson, et al., 2008), no images of real objects were used. Nonsense objects adapted from materials developed by Urbain et al. (2013) were used in place of the geometric figures used in previous studies. These abstract images were converted to grayscale to discourage the use of idiosyncratic verbal strategies that may rely on colour, and to better isolate the effect of the word forms themselves on word mapping. See Figure 2 for example images.

### Auditory Stimuli

Each experiment version (A/B) contained 12 target words that were evenly divided into three types that varied both on frequency and phonotactic probability: common words, rare words, and non-words. Each word type contained 4 bisyllabic nouns. Word types were chosen to maximize contrasts, and test whether words typical of Italian – that is, high in both frequency and phonotactic probability – would be mapped to objects faster than words that were low in both frequency and phonotactic probability. As described further below, common words were both frequently occurring and contained high phonotactic probabilities, and thus potentially would be more familiar sounding to the L2 exposure group. The rare words occurred only once throughout the 21 podcasts, and also had relatively low phonotactic probabilities. The non-words had extremely low phonotactic probabilities and were included to test whether sound combinations that violate typical Italian phonotactics would inhibit associative learning.

The frequency categorization and phonotactic probability (quantified as phonotactic score) calculations for each item were based on the Italian exposure podcasts. Transcripts for the podcasts were directly obtained from the source website, *News in Slow Italian.* The *spaCy* package, a natural language processing package in Python, was used to identify each word's grammatical category (e.g., noun, verb, article) and number of syllables (Honnibal & Montani, 2017). This information was used to compile a list of all bisyllabic nouns from the transcript, along with their frequency counts. Once all the bisyllabic nouns from the audio transcript were identified, they were sorted based on the total number of occurrences across the 21 podcasts. The 150 most frequent and 150 least frequent words were classified as high and low frequency respectively. High frequency words occurred between 7 and 252 times and low frequency words occurred only once. Non-words were created and the phonotactic scores of target words were determined based on a procedure carried out by another graduate student (see Appendix A - excerpt from Hoeppner 2024 thesis). Non-words were created by changing up to three phonemes from real words which resulted in highly improbable phoneme sequences. Phonotactic scores were calculated by generating a trigram model for the entire sequence of phonemes from the Italian exposure podcasts and then using the trained model to calculate log probabilities of the phonetic sequence of each test item. These scores were used as a measure of how prototypical the word structures were in Italian, based on the likelihood of their sound combinations.

Audio .wav files for each word were generated with Google cloud console's text-to-speech software using a female Italian voice. Both real words and non-words were checked by two native Italian speakers to ensure that they sounded natural, and that none of the non-words resembled existing Italian words. In addition, we also confirmed that none of the test words sounded like English words or were loan words from Italian that would be known to non-Italian speakers (e.g., "tempo", "grande", "solo"; see Appendix B for a complete list with frequencies and phonotactic scores of words and non-words used).

### 2.2.3 Continuous Listening Task

Podcasts for the testing session were extracted from the same series, *News in Slow Italian*, as the Italian listening period, but consisted of different podcasts not included as part of the listening period. For both experiment versions (A/B), we generated 20 unique audio files, each approximately 2 minutes in length that were played contiguously in approximately 10-minute segments. Audio files within each segment were from the same story and thus formed a coherent narrative. The four segments alternated between male and female speakers over the course of the listening task. One 2-minute audio clip from a segment later in the task was played prior to the first segment, presented continuously with the first segment. This repeated clip was introduced to help improve the signal to noise ratio (SNR) for our models, and as a validity measure, consistent with previous research (Crosse et al., 2021; Desai et al., 2021).

After the first 10 participants, a word detection task was introduced in order to ensure that participants were adequately attending to the task. Target words were selected for each listening segment on the basis of occurring multiple times throughout the segment. Version A contained 35 total target word occurrences and Version B contained 23 target occurrences. Each 10-minute segment contained an average of 7.25 (range: 5 to 10) instances of a target word.

## 2.3 Procedure

### 2.3.1 General Summary

An overview of the procedure is shown in Figure 1. Participant performance on identical measures was compared before and after a 3-week listening period. The pre-test and post-test experiments used counterbalanced sets of stimuli. Between the two testing sessions, participants listened to either natural speech in Italian (L2 group) or English (control group) in the form of daily podcasts.

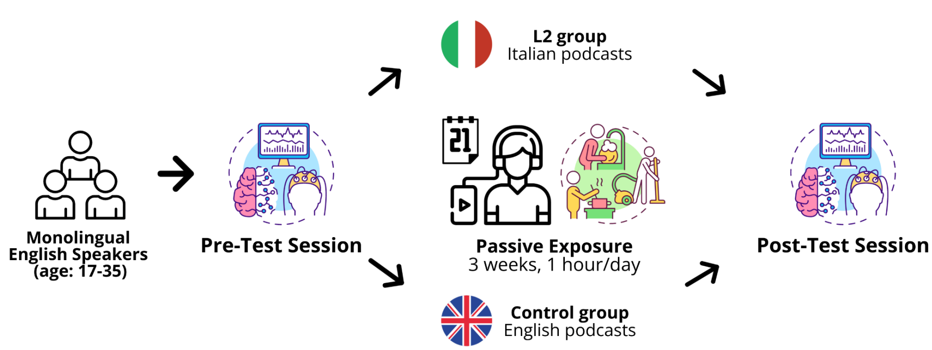


**Figure 1: Experiment design**

### 2.3.2 Session 1

After arriving at the lab, participants provided informed consent, and then filled out a brief Qualtrics survey asking about general demographic information and their language history while being set up for EEG. They then moved to a sound attenuated booth to complete three consecutive behavioural tasks while their EEG data were recorded. Each task had opportunities for breaks built in at regular intervals. In all experimental tasks, audio was played through external speakers at a comfortable volume and visual stimuli were presented on a 52 cm by 32 cm computer monitor placed approximately 75 cm in front of the participant. Responses were made through keypress of either the spacebar or number keys on a keyboard. EEG data were aligned with the auditory stimulus using trigger codes sent through the Cedrus Duo Stim tracker. Participants first completed the Word Rating task (which will not be discussed as a part of this thesis), followed by the Word Mapping task and the Continuous Listening task. After completing the three tasks, participants were given information and instructions about the 3-week listening period, and were then compensated for their participation in the first testing session.

### 2.3.3 Word Mapping Task (Behavioural Task)

The Word Mapping task, shown in Figure 2, was developed based on previous studies of word mapping (Graf Estes et al., 2007; Hay et al., 2011; Mirman, Magnuson, et al., 2008). This task was designed to compare the rate of learning object-label pairings across different word types. Each trial started with the auditory presentation of a test word. While the word was played, 500 ms after the start of the trial, two images appeared on either side of a fixation cross on the screen, and participants were asked to press 1 for left or 2 for right to indicate the appropriate image for the word. Immediately after they selected a response, feedback appeared on screen as the word “correct” or “incorrect” for 1 s. The task is designed such that participants must initially guess which of the images corresponded to the word that was played, but eventually acquire the correct associations through trial-and-error.

Prior to beginning the actual task, participants completed five practice trials, in which they heard English words corresponding to images of everyday objects. After each practice trial, they were given immediate feedback, ensuring they understood the task.

Experimental trials were divided into 11 blocks, with self-paced breaks between each block. Each block consisted of 24 trials, in which 12 words were presented twice in a randomized order. Each word was presented with its corresponding target image once on the left and the right in every block. The foil images consisted of images corresponding to other target words, also presented an equal number of times in each block.

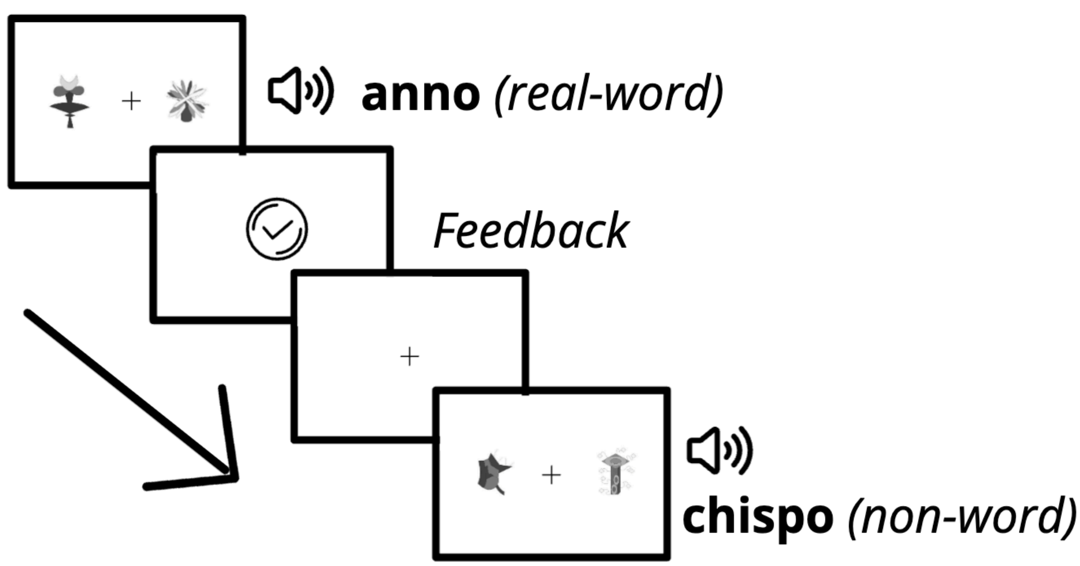


**Figure 2: Word Mapping Task**

### 2.3.4 Continuous Listening Task (EEG)

Participants were then asked to listen to approximately 40 minutes of Italian podcasts. To encourage attentive listening, participants were given a comprehension question every 10 minutes. They were asked: “What was spoken about in the last segment you listened to?” and given 3 multiple-choice options on the general topic (e.g., elections, a sporting event, etc.). Participants also completed a target word detection task (introduced after the first 10 participants). A word remained on screen during listening, varying between segments, and participants were asked to press the spacebar whenever they thought they heard the target word. This task was incorporated to increase task engagement, attention to the auditory stimuli and general alertness and motivation.

### 2.3.5 Listening Period

Participants were assigned to either the Italian listening (L2) group or the English listening (control) group based on a predetermined counterbalancing order. During the 21-day listening period, participants received daily emails with the link to the podcast they were meant to listen to and a survey to confirm they completed the task. The podcasts were embedded with ‘secret’ words that participants were asked to report in an online survey as a measure of protocol adherence. Each survey asked participants if they listened to the podcast for that day, and asked them to select up to six items, three to five of which were actually heard in the podcast, with the remainder serving as foil items.

Participants were encouraged to listen to the podcasts for an hour while doing other “everyday”, nonverbal tasks (e.g. housework, commuting, exercising). They were instructed to listen to each day’s podcast in a single uninterrupted session at any time within 24 hours from receiving their daily email. They were also explicitly told that to be eligible for the second testing session, they would be required to listen consistently and correctly report the multiple ‘secret’ words in the survey.

### 2.3.6 Session 2

Session 2 was almost identical to the first testing session. Participants filled out a survey asking about which activities they completed while listening to podcasts during the exposure phase. After being capped for EEG, they completed the three remaining experimental tasks (Word Rating task, Word Mapping task, and Continuous Listening task) using the alternate counterbalanced versions to those used in session 1. Following completion of the final task, participants were debriefed and compensated for their participation.

## 2.4 Data Analysis

### 2.4.1 Word Mapping Task

For each trial, each participant's image selection was scored as either correct or incorrect. Learning was steepest in the initial blocks of the experiment with mean accuracy reaching close to ceiling performance (88.2%) by block 3 (Figure 4). Therefore, each participant's mean accuracy was averaged across the initial three blocks of trials as a measure of initial learning and also across all 11 blocks as a performance measure across the entire experiment.

For both analyses (early learning and entire task), a 2x3x2 repeated-measures ANOVA was conducted on mean accuracy scores with group (control, L2) as a between-subjects factor and word type (common, rare, non-word) & session (1, 2) as within-subjects factors. If L2 exposure influences word mapping accuracy, we would expect to see a significantly stronger increase in accuracy from session 1 to session 2 in the L2 group, relative to the control group, as revealed by a group x session interaction. If L2 exposure has a selective benefit on mapping words more typical of the L2, we would also expect a significant interaction with word type such that all word types are learned equally at session 1 and by controls, but that at session 2 participants in the L2 group would have significantly higher accuracy for common words than for the rare words or non-words. Such an effect would be revealed by a significant word type x session x group interaction.

For two of the participants (one from each group), the Word Mapping task program unexpectedly malfunctioned, resulting in them only completing the first half of the trial blocks during one of the testing sessions. These two participants' data were excluded from the analyses of all 11 blocks, but retained in the analyses of initial learning.

### 2.4.2 Continuous Listening Task

**EEG Data Analysis**

The purpose of the EEG analysis was to test the impact of three weeks of L2 exposure on listeners' neural response to natural L2 speech. Using the mTRF toolbox (Crosse et al., 2016), neural data was mapped to two stimulus features: the acoustic envelope and word onsets.

***EEG Acquisition and Preprocessing***

EEG was recorded at a sampling rate of 512 Hz using a 64-channel Active-Two Biosemi system (Biosemi, Amsterdam), set up according to the 10/20 system. Additional electrodes were placed on the left and right mastoids, on the outer canthus of each eye, and below the left eye. Signals were recorded relative to the Common Mode Sensor active electrode and then re-referenced offline to the average of the left and right mastoid electrodes. DC offsets were maintained at ±20mV.

All EEG analyses were conducted using EEGLAB (Delorme & Makeig, 2004) and the mTRF Toolbox (Crosse et al., 2016). Data was converted to continuous neural data (CND) format as per CNSP guidelines (Di Liberto et al., 2023). In order to time lock the stimulus features to the EEG response, trigger codes were sent through the Cedrus StimTracker Duo as an indicator of when each audio file started. Neural data corresponding to each 2-minute audio file was extracted relative to a trigger code indicating the start of the file. The alignment was adjusted further by ensuring that audio onset timings (as detected by the StimTracker) corresponded to utterance onsets timings relative to the beginning of each file. Each participant's neural data underwent minimal preprocessing as is typical for continuous neural data analyses using the mTRF approach (Crosse et al., 2021). The segmented data was downsampled from 512 to 128 Hz, bandpass filtered from 0.5-15 Hz, and any bad channels were removed and replaced with a spline interpolation of all other channels. Bad channels were considered those with a standard deviation three times greater than that of the other channels.

### ***Stimulus Feature Extraction***

The auditory envelope of the podcast stimuli was extracted using the *mTRFenvelope* function from the mTRF toolbox (Crosse et al., 2016). This function computes the resampled temporal envelope of the audio signal using the Hilbert transform method providing a measure of the instantaneous signal amplitude. The root-mean-square (RMS) of the audio signal was computed over a window of samples and the resulting RMS intensity was raised to the power of 0.3 to model human hearing (Crosse et al., 2021; Lalor & Foxe, 2010). The envelope was downsampled to 128 Hz and then normalized and trimmed to match the length of the corresponding neural data for each trial.

The word onsets were extracted using a 2-step process. First word onset timings were extracted using the BAS WebMAUS aligner set to Italian (Kisler et al., 2017), by aligning the individual audio transcripts (taken from the *News in Slow Italian* website) and audio files for each 2-minute trial. The BAS WebMAUS aligner is a web service that aligns the phonological transcript to the speech signal and outputs phonetic segmentation and labeling. The word onset timings were then converted to a stimulus vector with a sampling rate of 128 Hz consisting of zeros and ones (with one indicating the start of a word), corresponding to the length of the audio file.

### ***Removal of Bad Trials***

Before modeling the data, any 2-minute trials with large EEG artifacts (e.g., out of range values) that impacted all channels were removed using a simplified cross-validation procedure, as follows. The neural response to the envelope was modelled using the *mTRFcrossval* function (Crosse et al., 2016) with no regularization ($\lambda = 0$). Error values were generated for the model fit to each trial at each electrode and then averaged across all channels. Trials with outliers – prediction errors greater than 3 standard deviations away from the mean – were removed at the participant level prior to further analysis. For each participant, no more than 6 trials were removed ($M = 1.13$, $SD = 1.52$).

### *Fitting Encoding Models*

Temporal response functions (TRFs) were generated in the forward modeling direction, predicting the neural response from the stimulus. The TRF is a regularized ridge regression that maps stimulus features to each recording channel over a range of time lags. The regression weights reflect the strength of the stimulus-EEG mapping at individual time latencies (Di Liberto et al., 2021). In order to best assess the auditory response, the time window was set from -100 ms to 500 ms. The TRF was optimized using ridge regression to prevent overfitting of the stimulus to data-specific noise. The optimal lambda value (ridge parameter) was determined using a leave-one-out cross-validation approach. The best lambda value in the range of $10^{-6}$ to $10^{6}$ was then used to train the forward models using the *mTRFtrain* function and obtain model parameters and prediction correlations (Crosse et al., 2016).

These models were generated using all trials at the participant level, producing prediction correlations generated for each electrode, which were then averaged together to give a model fit at the individual level. Results were combined across both stimulus versions (A/B) and all participants at the group level to give average model fit for three different models: (1) the acoustic envelope, (2) the word onsets and (3) a combined model including both the envelope and the word onsets as predictive features.

### *Analyzing Models*

These three models were tested to see whether each one could predict the neural response significantly above chance (t-test compared to zero) across all subjects and sessions. Next, a repeated measures ANOVA with group (control, L2) as a between-subjects factor and session (1, 2) as a within-subjects factor was conducted for each of the models. These comparisons were made to test the impact of L2 exposure on lexical segmentation in L2 speech. We would expect reliable tracking of the acoustic envelope by all participants across both sessions (model 1), but a significant session x group interaction for both the word onset model (2) and the combined model (3), with these models predicting the neural response better for participants in the L2 group post-exposure. Such an effect would show up in a significantly higher prediction correlation of the word onset and combined models at session 2 for the L2 group as compared to session 1 and controls. In order to test whether word onsets have a unique contribution to the neural representation beyond the acoustic envelope alone, we subtracted the mean prediction correlation of the envelope model (1) from the combined model (3) for each participant. If our hypothesis that L2 exposure facilitates word segmentation is supported, we would expect a significant group by session interaction such that the prediction correlations for this difference model (4) would only improve for the L2 group at their second testing session.

**Behavioural Data Analysis**

For each participant, performance on the four multiple-choice questions was computed as the proportion of questions answered correctly. Performance was then assessed using a repeated measures ANOVA with condition (control group, L2 group) as a between-subjects factor and session (1, 2) as a within-subjects factor.

Target detection data was available for 34 participants. Hit rate was calculated as the number of times the spacebar was pressed within 2 seconds of a target word divided by the total target occurrences. Descriptive statistics are reported for the overall hit rate and number of false alarms.

# Chapter 3

# 3 Results

## 3.1 Podcast Exposure Period

The three most frequently reported activities that participants engaged in while listening to the podcasts were cleaning (74%), commuting (74%) and cooking/eating (65%). All reported activities are summarized in Appendix C.

Participants in the control group reported listening to an average of 20.4 podcasts (range: 17 - 21) out of which they correctly identified 89.9% of hidden words (range: 64% - 99%) and incorrectly selected 1.25% of distractor items (range: 0 - 9.5%). Participants in the L2 group reported listening to an average of 20.7 podcasts (range: 18 - 21) out of which they correctly identified 88.1% of hidden words (range: 52.9% - 100%) and incorrectly selected 2.64% of distractor items (range: 0 - 29.4%). These results confirm that participants in both groups showed good compliance with the experimental protocol.

## 3.2 Word Mapping task

### 3.2.1 Overall Analysis

In contrast to our hypotheses, we did not find a significant 3-way interaction between group, session, and word type, indicating that the two groups did not differ significantly in word-mapping rates between the three types of words from session 1 to session 2 ($F(2,84) = 0.044$, $p = .957$; Figure 3). In addition, we also did not find a significant 2-way interaction between group and session ($F(1,42) = 0.451$, $p = .506$), indicating that any change in performance from session 1 to session 2 was similar between the two groups. We did not find any main effects of session ($F(1,42) = 1.052$, $p = .311$) or word type ($F(1,42) = 2.139$, $p = .124$), nor did we find a main effect of group ($F(1,42) = 1.378$, $p = .247$).

A marginally significant session x word type interaction ($F(2,84) = 2.619$, $p = .079$; Figure 4) was revealed, indicating that across both groups, the word types were learned differently between sessions. Although only marginally significant, we ran post-hoc contrasts using the *emmeans* package in R to better understand this trend. Contrasts showed that there were no significant differences between the word types at session 1 (all $p > .294$). In comparison, at session 2, there was a significant difference in accuracy between common and rare words ($t(162) = 2.602$, $p = .027$), with common words being mapped more accurately. There were no significant differences between common words and non-words ($t(162) = 0.973$, $p = .595$) or between rare words and non-words ($t(162) = -1.629$, $p = .236$).

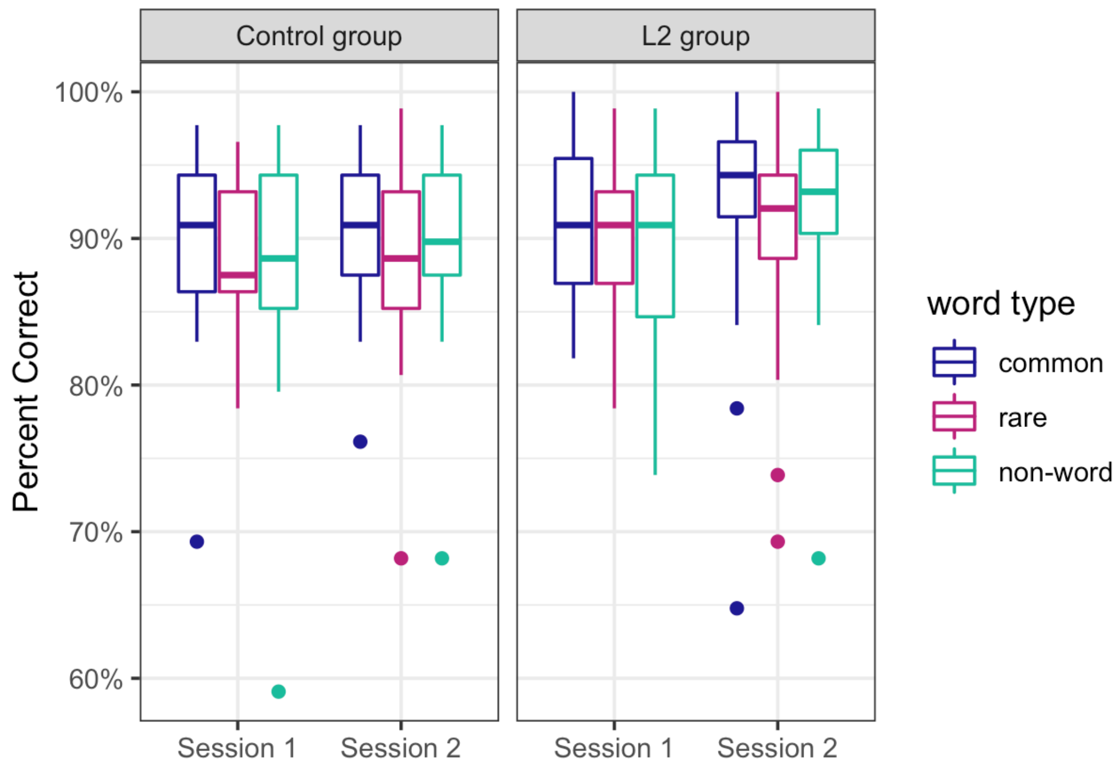


**Figure 3: Box plot of word mapping accuracy across the entire experiment by group, session and word type.**

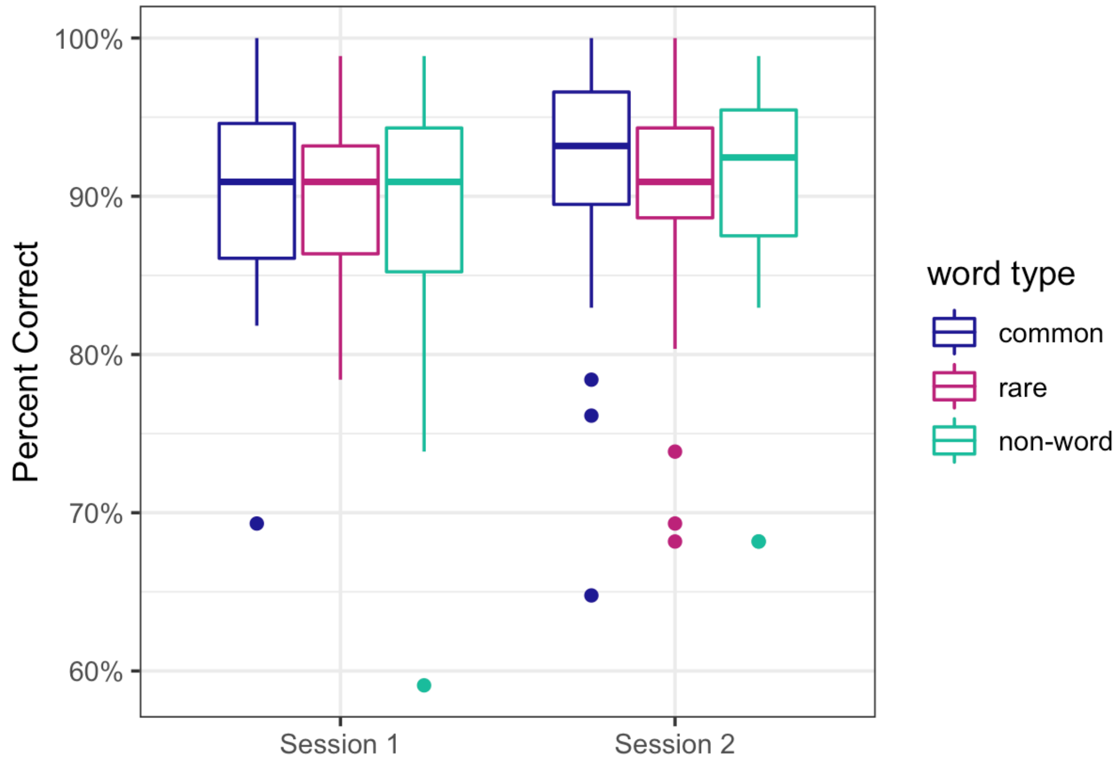


**Figure 4: Box plot of word mapping accuracy over all blocks by word type and session, across both groups.**

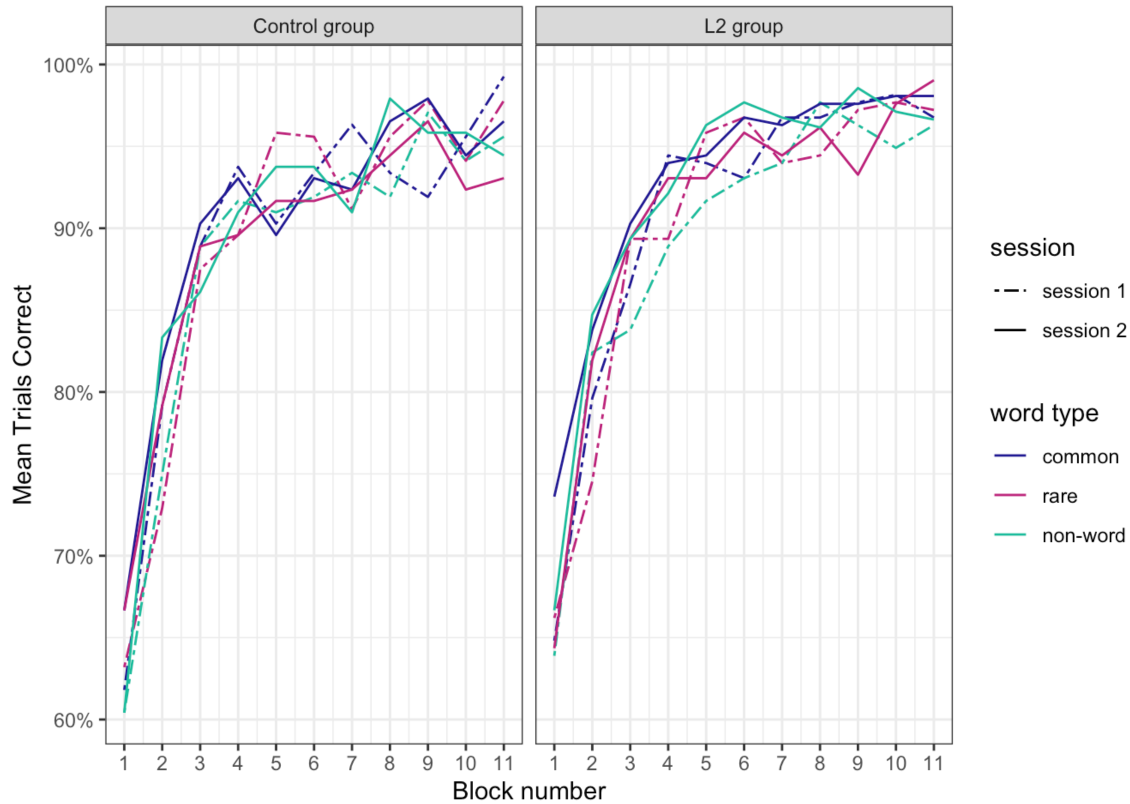


**Figure 5: Line graph showing the mean word mapping accuracy at each block by group, session and word type.**

### 3.2.2 Early Learning Analysis

Contrary to our hypotheses, we did not find a significant 3-way interaction between group, session, and word type ($F(2,86) = 0.387$, $p = .680$; Figure 6), nor did we find a group x session interaction ($F(1,43) = 0.061$, $p = .805$). These results indicate that both groups showed a similar level of change from the first to the second session across all word types.

As in the overall analysis, we again did not find a main effect of group ($F(1,43) = 0.567$, $p = .456$) or word type ($F(1,43) = 1.354$, $p = .264$). However, the mean accuracy over the first three blocks of trials differed between sessions (Effects of Session: $F(1,43) = 4.089$, $p = .049$; Figure 7). Post-hoc contrasts showed that, overall, across word types and groups, accuracy at session 2 (*M* = 79.5%, *SD* = 11.5%) was significantly greater than accuracy at session 1 (*M* = 76.2%, *SD* = 12.5%; *t*(134) = -2.69, *p* = .008).

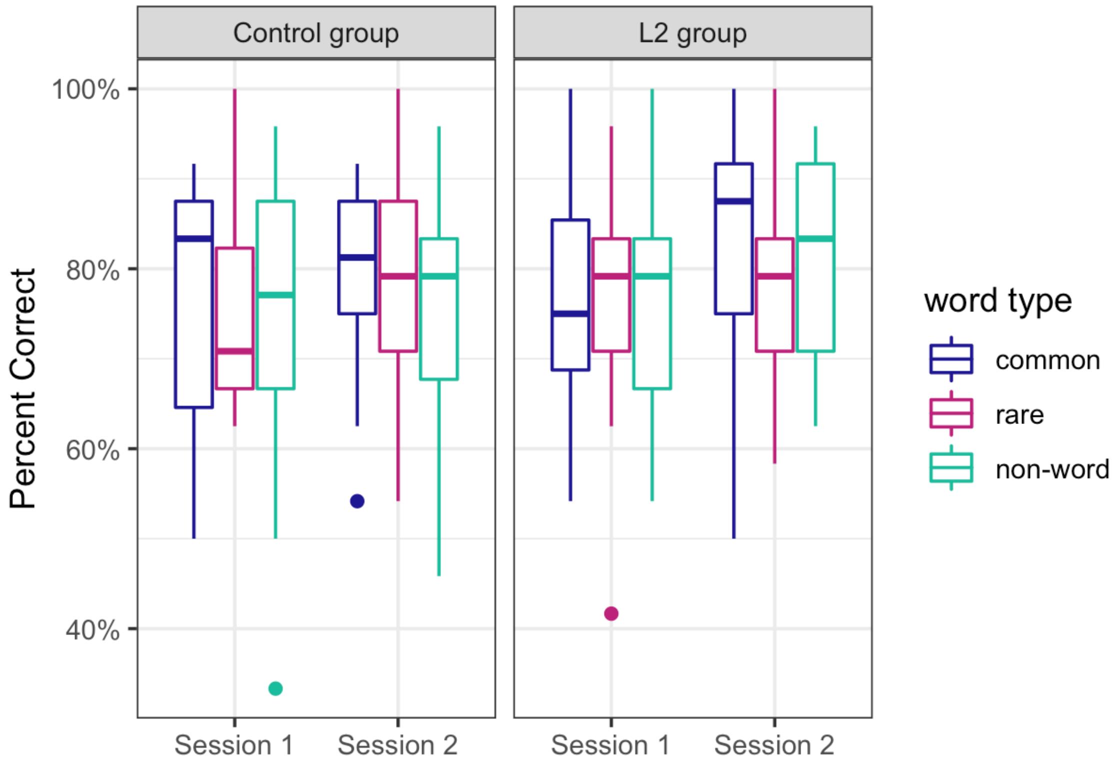


**Figure 6: Boxplot of word mapping accuracy over the first three blocks by group, session and word type.**

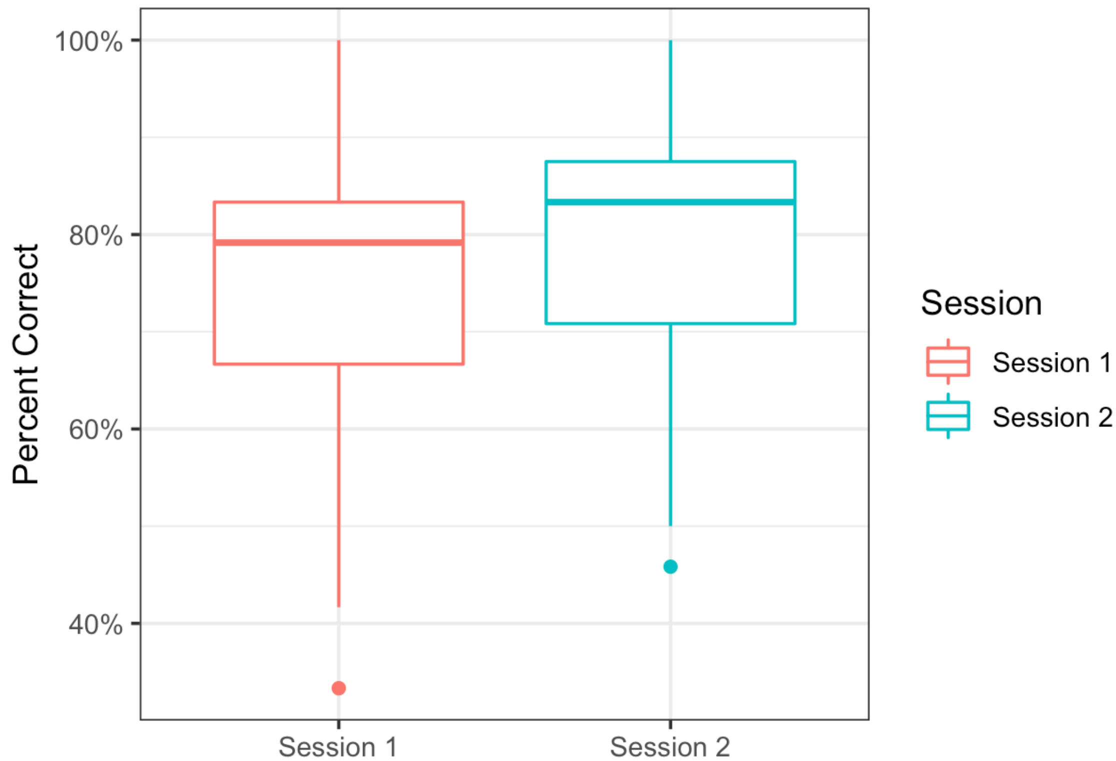


**Figure 7: Boxplot of word mapping accuracy for the first three blocks by session, across all participants and word types.**

## 3.3 Continuous Listening Task

### 3.3.1 EEG Results

The prediction correlations for each of the tested models are shown in Figure 8. The corresponding TRFs and topoplots are shown in Figure 9.

**Acoustic Envelope**

The prediction correlations for the acoustic envelope model ($M$ = 0.048, $SD$ = 0.020) were significantly greater than zero ($t(87)$ = 22.23, $p < .001$), indicating that this feature can reliably predict the neural response to continuous L2 speech across all participants and both sessions.

The acoustic envelope predicted the neural response equally well across both sessions ($F(1,42) = 2.89$, $p = .096$) and groups ($F(1,42) = 1.55$, $p = .220$). We did not find a significant group x session interaction ($F(1,42) = 0.644$, $p = .427$), indicating that acoustic envelope tracking did not change as a function of L2 exposure.

**Word Onsets**

The prediction correlations for the word onset model ($M = 0.023$, $SD = 0.014$) were significantly greater than zero ($t(87) = 15.13$, $p < .001$), indicating that this feature can reliably predict the neural response to continuous L2 speech in all participants and across both sessions.

Contrary to our hypothesis, prediction correlations for the word onset model did not differ significantly between sessions ($F(1,42) = 1.63$, $p = .209$), nor was there a significant group by session interaction ($F(1,42) = 0.712$, $p = .404$). We also did not find any group differences ($F(1,42) = 1.52$, $p = .224$). These findings indicate that the neural response to word onsets did not improve after L2 exposure, and word onsets predicted the neural response of all participants equally well at both sessions.

Based on prior literature we would expect the acoustic envelope, which is reliably tracked regardless of attention or linguistic proficiency, to capture more sound features, and therefore give higher prediction correlations than the word onset model (Gillis et al., 2021; Zou et al., 2019). As expected, the acoustic envelope model predicted the neural response better than the word onset model ($M_{diff} = 0.0253$; $t(87) = 21.86$, $p < .001$).

**Combined Model (Acoustic Envelope + Word Onsets)**

The prediction correlations for the combined model ($M = 0.049$, $SD = 0.021$) were significantly greater than zero ($t(87) = 21.81$, $p < .001$), which is expected given that the two features that went into it, the acoustic envelope and word onsets, were both tracked reliably by all participants across both sessions.

The combined acoustic envelope and word onset model also predicted the neural response equally well across both sessions ($F(1,42) = 2.720$, $p = .107$) and groups ($F(1,42) = 1.513$, $p = .226$). We did not find any significant session by group interactions ($F(1,42) = 0.444$, $p = .509$).

**Difference Model (Combined - Acoustic Envelope)**

The prediction correlations of the combined model were significantly greater than the envelope only model ($t(87) = 6.0$, $p < .001$), indicating that across both sessions and groups, word onsets predicted the neural response significantly beyond any variability accounted for by the acoustic envelope.

Given that the combined and envelope model did not show any effects of group or session, we did not expect the prediction correlations of the difference model to vary as a function of the L2 exposure. Nonetheless, we confirmed this idea by calculating the difference between models at the participant level and conducting a repeated measures ANOVA on the differences in prediction correlations. Comparing the difference between the combined model and acoustic envelope model showed that word onsets uniquely predicted the neural response equally across sessions ($F(1,42) = 0.196$, $p = .660$) and groups ($F(1,42) = 0.208$, $p = .650$), and did not show any significant session by group interactions ($F(1,42) = 0.250$, $p = .619$). This indicates that, counter to our original hypothesis, participants in the L2 group did not show an increased neural response to word onsets in the L2.

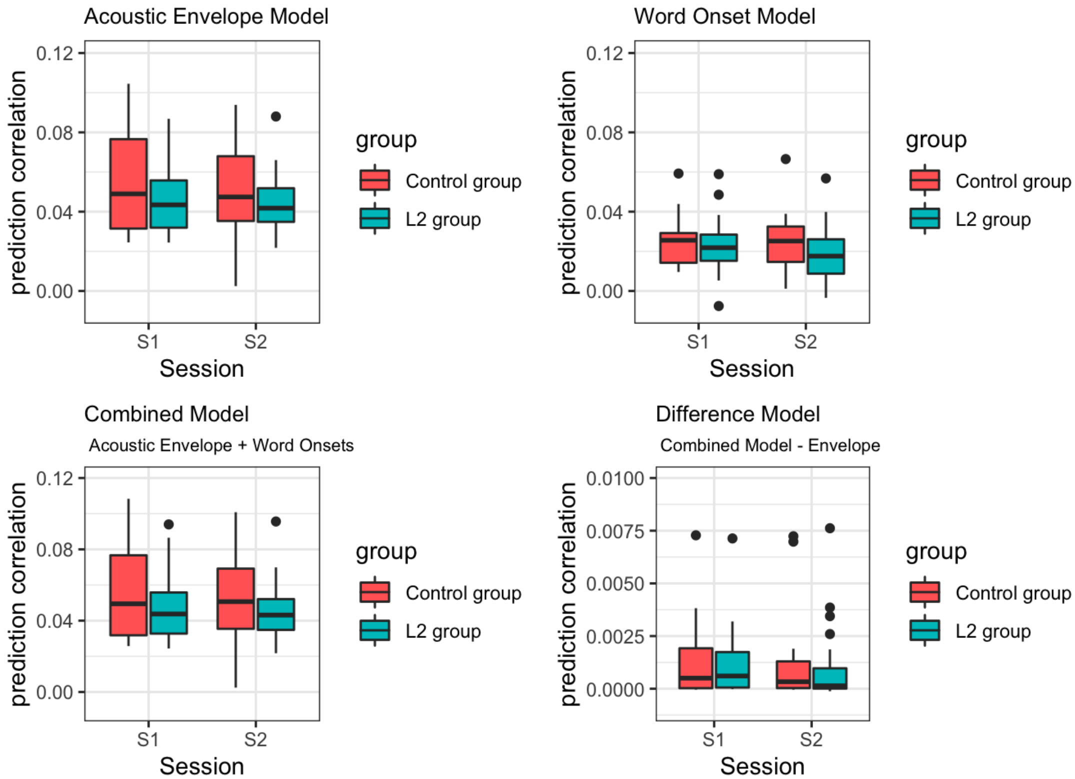


**Figure 8: Boxplots comparing the mean prediction correlation for each model as a function of group and session.**

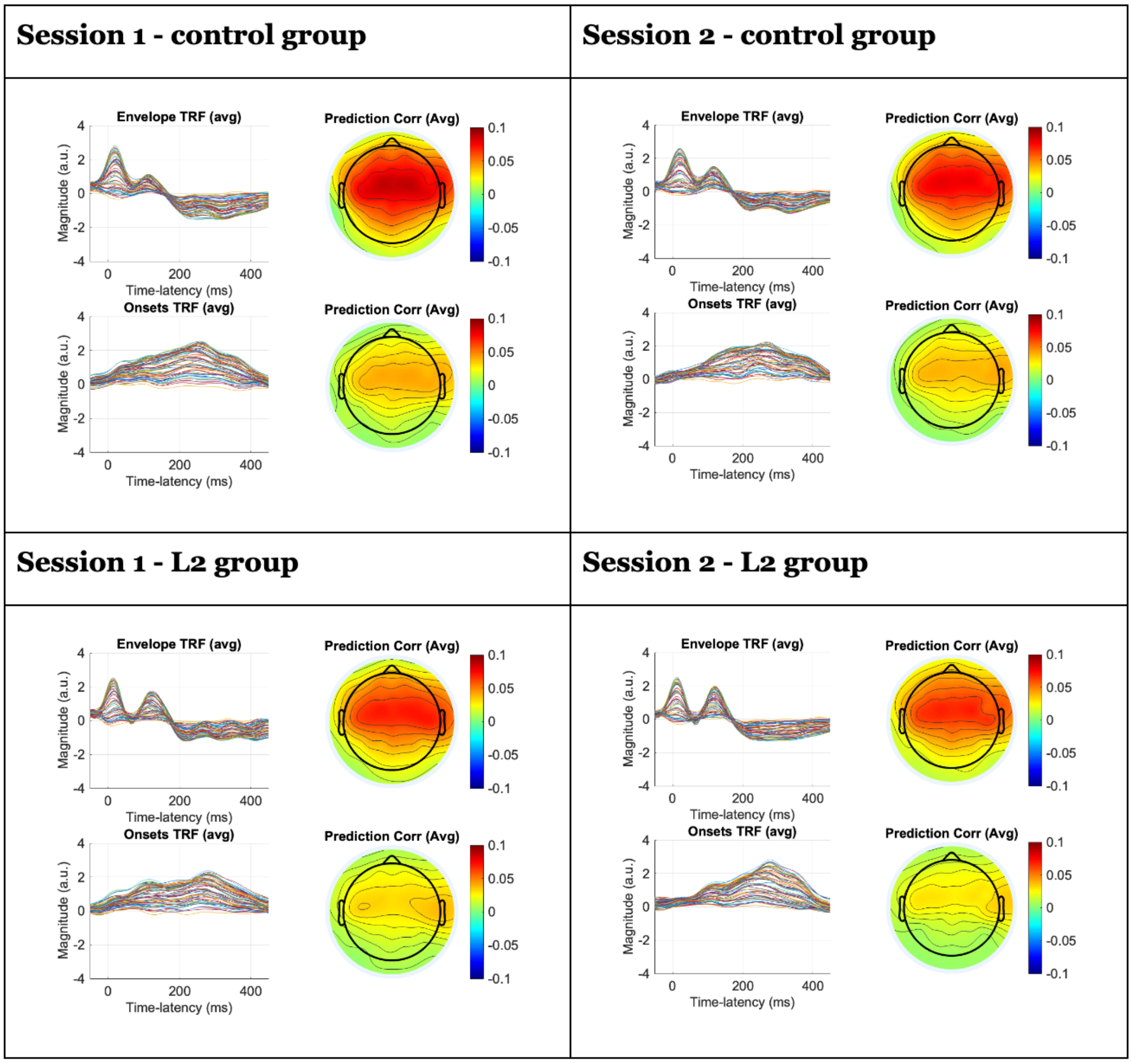


**Figure 9: TRFs and topoplots**

*For each group and session, the top and bottom rows of the figure show the acoustic envelope model (1) and word onset model (2) respectively. All TRF weights and prediction correlations were averaged across participants for each session and group. Scalps maps plot the prediction correlations at each channel.*

### 3.3.2 Behavioural Results

Participants performed reasonably well on the comprehension questions, answering an average of 75.2% of the 3AFC questions accurately (*SD* = 22.5%; chance is ~33%). An ANOVA comparing question accuracy found that accuracy on this measure was not significantly impacted by group, ($F(1,42) = 0.258$, $p = .614$), testing session ($F(1,42) = 0.947$, $p = .336$), nor their interaction ($F(1,42) = 0.001$, $p = .981$).

For the target detection component of the Continuous Listening task, participants performed relatively poorly, with an overall hit rate of 33.4% (range: 0 - 81%). This relatively poor performance may be attributed to the difficulty of the task. False alarms, defined as any click that did not occur within 2 s of a target word, occurred an average of 15.9 times (range: 0 - 53) throughout the 40-minute task.

# Chapter 4

# 4 Discussion

The purpose of this study was to test the extent to which statistical learning mechanisms extend to real-world language acquisition through a passive L2 exposure paradigm. We hypothesized that three weeks of passive L2 exposure would enable adult listeners to become sensitive to phonotactic patterns that further aid in word segmentation and subsequent word mapping. To investigate this question, we tested participants on two different tasks. The first task was the Word Mapping task, which provided a behavioural measure that allowed us to compare the rates at which participants could learn object-label pairings for words that varied in frequency and phonotactic probabilities. The second task was the Continuous Listening task, which provided an EEG measure of the neural response specific to word onsets in natural L2 speech. As phonotactic probabilities are known to play a role in both word segmentation and word mapping ability, these two measures together provide an indicator of learners' sensitivity to phonotactic structure. We were interested in quantifying any changes in these measures that take place over time as a result of passive L2 exposure. Results of the two measures are discussed in turn in the following sections.

## 4.1 Word Mapping Task

The Word Mapping task was designed to test whether passive exposure to an L2 facilitates subsequent vocabulary mapping of words that are consistent with L2-specific phonotactic probabilities. While there are numerous important aspects of word learning, word mapping is an essential step in the acquisition of a new language (Wojcik et al., 2022). Based on previous studies of word mapping (Graf Estes et al., 2007; Hay et al., 2011; Mirman, Magnuson, et al., 2008), we expected acquisition of L2 phonotactics to facilitate learning object-label associations for L2 consistent words and inhibit learning object-label associations for L2 inconsistent words. By testing three word types – common words with both high frequency and high phonotactic probabilities, rare words with both low frequency and low phonotactic probabilities and non-words with extremely low phonotactic probabilities – we expected that participants familiar with L2 word structures would learn common words faster than rare or non-words and thus also perform more accurately on the task as a whole.

In contrast to our hypothesis, results from the task did not show the expected effects of the experimental manipulation, with participants in both the L2 and control group performing equally well across all word types, and showing similar improvements from session 1 to session 2. This indicates that three weeks of passive L2 exposure did not provide a measurable advantage to the L2 group for mapping L2 consistent words to object referents. Across all experimental blocks, there was a marginally significant interaction between session and word type. Participants in both the L2 group and the control group were more accurate at mapping common words than rare words in the second session. While this is an interesting numerical trend, it may not reflect a meaningful difference given the marginal interaction. Collecting the full sample will provide additional power to evaluate the robustness of this finding and potentially reveal a group interaction . We did not anticipate the differential mapping of word types in our control group, but it is possible that the 40 minutes of L2 exposure from the first testing session was sufficient to induce generalizable knowledge of L2 phonotactic regularities, leading to an overall session 2 benefit in mapping the common words. Previous studies have shown that as little as 7 minutes of L2 exposure can lead to some phonotactic knowledge (Gullberg et al., 2010; Kittleson et al., 2010). In addition, analyses of the initial learning blocks found that participants in both groups demonstrated an improvement from session 1 to session 2 across all word types. This improvement may be attributed to general task practice effects, as the initial session provided an opportunity for participants to become more skilled at the task overall.

There are a number of possible reasons why the results of the present study do not show conclusive evidence that L2 exposure leads to an advantage in learning L2 consistent object-labels. One key difference between the current task and previous experiments using this paradigm (Hay et al., 2011; Mirman, Magnuson, et al., 2008) is that the word types in the current study differ on phonotactic probabilities, rather than transitional probabilities. While both measures are highly related and make similar use of the distributional cues in the language input, even the rare words with a low phonotactic score are still valid words and do not explicitly violate expectations for what a word should be in the same way that 'part-words' used in past studies do. Therefore, the inhibition of word mapping for words with low transitional probabilities may not extend to those with low phonotactic probabilities, leading to similar word mapping rates across word types despite L2 exposure.

Another explanation is that adults may be better equipped to process and subsequently map words with lower phonotactics compared to infants. Research has shown that by 2 years of age, toddlers are able to learn words with high transitional probabilities, low transitional probabilities and even words that violate low transitional probabilities equally well. They only show a disadvantage for words that violate high transitional probabilities (Lany et al., 2024). In addition, a study looking specifically at the impact of phonotactic probabilities found that they becomes less influential as infants develop; while younger infants are only able to learn words consistent with the phonotactics of their native language, older infants and infants with larger vocabularies are able to learn words with lower phonotactic probabilities (Gonzalez-Gomez et al., 2013). This leads to the possibility that the impact of phonotactics on learning rate in adult participants may be too subtle to be detected by the present accuracy measure.

The non-words used in our study, while they had highly unlikely sound combinations for the Italian language, could possibly have been treated as neutral non-words for native English speakers. Despite being designed to have a very low phonotactic score and to theoretically inhibit subsequent mapping, they may have possibly been so different from words that they were treated neutrally. Previous word mapping studies using artificial languages have shown that words that violate specific expectations (part-words consisting of syllables spanning word boundaries) are difficult to map, but other words made up of entirely new sounds (syllables that did not occur in the speech stream) can be mapped just as easily as 'true' words with high transitional probabilities (Graf Estes et al., 2007; Karaman et al., 2024; Lany et al., 2024; Mirman, Magnuson, et al., 2008). In the present study, the phoneme substitution method used to generate non-words may have introduced unfamiliar syllables that were no longer constrained by the same expectancy effects as real words with low transitional probabilities. To better understand whether the non-words had the intended perception, we would also need to know how likely their sound combinations were for participants' native language, English. Although the sound combinations were unlikely in our test L2 (Italian), it is well known that one's native language has a strong influence on language processing (Brodbeck et al., 2024). Another way to better understand this measure would be to compare the word mapping rates for the different word types among native Italian speakers. Unlike our current participants, Italian speakers would already be familiar with even the 'rare' Italian words so we would not expect them to differ in their ability to map common versus rare words. In contrast, we would expect Italian speakers to map words with inconsistent phonotactics (non-words) less well than words that are already a part of their vocabulary.

### 4.1.1 Limitations and Future Directions

An important factor is that we have not yet met the intended sample size of 40 participants in the L2 group and 30 participants in the control group for the project. With the various experiment versions and individual differences in word learning ability, the power provided by our current sample was inadequate to reveal significant effects of our key experimental manipulation. While the effects are not significant, there are numerical differences in how participants in the L2 group learned common versus rare words, as shown in Figure 3 and Figure 6. Recruiting additional participants will allow us to determine whether these numerical trends reflect true yet small effects or the lack of any reliable differences.

A limitation of the current study design is that we cannot distinguish the effect of frequency from the effect of differences in phonotactic probability. Although the word types for the Word Mapping task were selected to maximize any effects of L2 exposure, there were no non-words with high phonotactics. Including these would have helped clearly show the effects of the manipulation and disentangle the effects of familiarity with specific words from more general learning of sound patterns. By intentionally selecting words high/low in both frequency and phonotactic probability we limit our understanding of which factor(s) are driving changes in accuracy. The impact of both word frequency and phonotactic probabilities on word mapping ability could potentially be explored further by conducting the same study on native speakers to see how they process the different word types. We could also further assess the neural response to these different word types by deriving ERPs from the current data.

Another possibility is that our task was too easy for our participants, with many participants reaching ceiling performance quite quickly. If the reason our measure did not show a specific effect of word types is that adults are too good at the Word Mapping task, increasing task complexity might lead to slower learning and subsequently highlight differences in the 'learnability' of specific word types. Increasing the number of total words to learn or drawing distractor items from a separate stimulus set may potentially lead participants to take longer to reach ceiling performance.

Finally, there are other analysis methods that might be more sensitive to subtle differences in the learning trajectory. Once our full sample is collected, data from this task could be further analyzed using logistic regression based growth curves, as developed by Mirman and colleagues (Mirman, Dixon, et al., 2008) to analyze data for this specific type of task.

## 4.2 Continuous Listening Task

The Continuous Listening task was conducted to test whether three weeks of passive L2 exposure changes participants' neural response to natural L2 speech. We specifically compared how well an acoustic envelope model, a word onset model, and a combined (acoustic envelope + word onset) model would predict the neural response before and after exposure. Overall, we found that the ability of these stimulus features to predict the neural response was not enhanced by L2 exposure. We also did not see any significant changes between the two testing sessions.

For the acoustic envelope model, we expected all participants to successfully track the envelope, which reflects bottom up sensory level processing (Brodbeck et al., 2024; Di Liberto et al., 2021; Gillis et al., 2021). Although the spatio-temporal profiles of the TRF may shift depending on language proficiency, the acoustic envelope can still be reliably tracked by L2 listeners (Di Liberto et al., 2021; Zou et al., 2019). Results show that all participants were sensitive to low level acoustic features of the speech signal, as revealed by prediction correlations for the acoustic envelope model that were significantly greater than zero.

The word onset model was used as a measure of lexical segmentation as per Brodbeck and colleagues (2018). Since word onsets often coincide with periods of no sound to sound, as is the case at utterance boundaries, we would expect word onsets to reliably predict the neural response to a certain extent, when considering word onsets as an independent feature (i.e., when the acoustic envelope is not included in the model). Therefore, it is not surprising that for all participants at both session 1 and session 2, the word onsets predict the neural response significantly above chance. However, in this word onset only model, the response to word onsets in this model may reflect acoustic properties that are correlated with word onsets, rather than "true" sensitivity to word boundaries per se.

To determine the *unique* contribution of word onsets, we used the difference between the combined model and the acoustic envelope to factor out the variance accounted for by the acoustic envelope, thereby better capturing sensitivity to word onsets that occur in the absence of salient acoustic cues. We had hypothesized that participants with no prior knowledge of or exposure to Italian would not be able to detect word onsets in continuous speech beyond the acoustic changes that would be reflected in the acoustic envelope. However, contrary to our expectations, even at session 1, the combined model performed better than the model with only the acoustic envelope. This result suggests that all participants showed some 'baseline' sensitivity to word onsets in L2 speech. We did not find any interactions with group or session for these difference values, indicating that the contribution of word onsets beyond that of the envelope did not change significantly in response to exposure, in contrast to our main experimental hypothesis.

If, as our results suggest, participants are able to reliably track word onsets at baseline, this may be due to several reasons. For one, over the course of 40 minutes of listening, the participants are no longer completely naïve listeners. It is possible that relatively brief exposure to L2 speech is sufficient to support some learning of language-specific cues to word boundaries, in line with previous studies that have shown sensitivity to language patterns after as little as 7 minutes (Gullberg et al., 2010; Kittleson et al., 2010). If the tracking of word onsets in session 1 was driven by L2 exposure during the earlier parts of the listening task, this could potentially be tested by comparing model fit for the initial trials to the later trials at the group level. This may reveal changes in neural tracking over the 40 minutes of exposure. Alternatively, there is also a possibility that even before the initial 40 minutes of exposure, English speakers are sensitive to word onsets in Italian. This may be facilitated by a combination of congruent language-specific strategies used for L1 segmentation (i.e. overlap between the L1 and L2) and universal language general cues that can be applied to L2 input (Gilbert et al., 2021). Language general cues include various overlapping segmentation cues such as stress patterns, utterance boundaries and cognates or other anchor words found in natural speech (Cunillera et al., 2010; Sahni et al., 2010; Sohail & Johnson, 2016).

Research has shown that in order to effectively model the neural response to linguistic features, participants must be actively attending to it (Brodbeck et al., 2018; Crosse et al., 2021; Vanthornhout et al., 2019). To encourage attention, we included comprehension questions as a behavioural measure, in which participants had to report what the segment was about. Participants in both groups showed above chance performance on this task, potentially by making use of cognates or loan words in the speech in order to determine the overall content. Results for the target detection task show that although participants did detect some words (33.4%), they missed many as well. Despite the low hit rate and relatively high false alarm rate for this task, the task achieved its purpose of keeping participants on task and giving them something to focus on.

### 4.2.1 Limitations and Future Directions

It is possible that the current analyses may not show any changes in the word onset response because we averaged across all scalp channels across a relatively large time window. From prior studies (Brodbeck et al., 2018; Karunathilake et al., 2023), we know that the response to word segmentation is typically found around 100 ms post stimulus onset, so limiting our time lags to a smaller, a priori defined window may help more closely align the EEG response to word segmentation specifically. In future analyses, selecting a subset of electrodes (e.g., frontocentral channels that typically show the largest responses to auditory and speech processing) and reducing the maximum time lag may be a more sensitive approach to investigate potential differences as a function of our experimental manipulation.

In the current thesis, the only measure of word segmentation that we analyzed was word onsets. However, there may be other lexical or sublexical features that are better able to capture any changes in language processing that take place during early stages of language acquisition. Future research could further investigate the temporal response function to other stimulus features such as phoneme-surprisal, word-surprisal, and different word classes (e.g., content words vs function words). As the previous study by Alexander and colleagues (2023) showed, passive L2 exposure for 2 weeks can enable some generalized word form knowledge, as assessed through an explicit behavioural task. Looking at these additional features would allow for a more holistic assessment of any changes at the neural level that may accompany these behavioural effects. In addition to assessing the impact of other stimulus features, it may also be useful to generate more complex models that simultaneously account for other acoustic features such as phoneme or syllable onsets. These would allow us to more definitively isolate any unique contribution of word onsets to the neural response to continuous speech.

In both the Word Mapping and Continuous Listening tasks, the lack of significant group differences could also be attributable to the limited and unbalanced sample size. We found that more participants dropped out of the control group ($n$ = 9/27; 33%) than the L2 group ($n$ = 5/32; 15.6%). This may be due to the additional motivation that comes with listening to or learning a new language. Anecdotally, participants in the control condition found listening to the content-matched English podcasts to be monotonous. The consequence is that the participants that completed the study – listening consistently for the full three weeks – were likely more motivated than those that dropped out, as the subjective experience of listening to English podcasts was relatively unpleasant. Future studies could address this issue by introducing more variable podcast content. In the current study, by recruiting an additional 25 participants to reach our intended sample size of 70, we may compensate for the selective attrition of control participants and achieve a more representative sample across both groups.

## 4.3 General Discussion and Conclusions

Across the two tasks, we did not find any significant effects of the 3-week L2 exposure, but we did find indicators that participants in both groups have some sensitivity to the phonotactic regularities of the L2. Participants in both the control group and L2 exposure group were sensitive to word boundaries across both sessions, and at the second testing session, showed improved word mapping ability for test items with high phonotactics as compared to those with low phonotactics. As none of our participants received any form of training or explicit L2 instruction, this improvement in word mapping for common L2 words provides some evidence that statistical learning mechanisms can extend to real world language acquisition.

In the Word Mapping task, equal performance across word types at session 1 was followed by higher word mapping accuracy for common versus rare words at session 2, providing evidence of phonotactic knowledge acquisition in both the L2 and control group. This improvement could potentially be due to the ~40 minutes of L2 exposure presented in the first testing session during the Continuous Listening task, despite the 3-week delay. However, as the interaction between word type and session was only marginally significant, it is also possible that these changes can be explained by some pre-existing knowledge or shared linguistic features between English and Italian. If we find a significant interaction between word type and session even after reaching our intended sample size of 70 participants, we would be able to rule out any baseline phonotactic knowledge of Italian. On the Continuous Listening task, if there was rapid learning of phonotactic regularities and acquisition of word boundaries within the initial listening segments, this may have been masked by averaging across the relatively long duration of the audio stimuli to generate the encoding models. This might explain why contrary to our expectations, we found that word onsets could reliably predict the neural response to continuous speech even at session 1. The current pattern of findings suggests that the 40 minutes of L2 exposure from session 1 was sufficient for participants to become sensitive to some L2 phonotactic regularities, and to maintain this knowledge over a 3-week delay.

As there is evidence of phonotactic knowledge being acquired even before the start of the 3-week listening period, it is possible that any further changes in speech processing induced by L2 exposure would be undetectable by the current measures. In addition, as just previously mentioned, the interaction between session and word type for the word mapping task was only marginally significant. Therefore, a plausible alternative explanation is that English speakers have some sensitivity to word boundaries in Italian without any prior exposure. In order to test whether learning occurred during the first testing session, we could potentially run another control group of participants who do not receive any L2 exposure at session 1.

Although the current investigation did not result in any conclusive evidence that passive language exposure facilitates word segmentation, there may still be differences in other aspects of word learning that are not captured by tracking word onsets. Further testing may show improvements in word mapping ability after exposure or changes in the processing of other linguistic features which would make a clearer case for the benefits of ambient L2 exposure.

# References


Alexander, E., Van Hedger, S. C., & Batterink, L. J. (2023). Learning words without trying: Daily second language podcasts support word-form learning in adults. *Psychonomic Bulletin & Review*, *30*(2), 751–762. https://doi.org/10.3758/s13423-022-02190-1

Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, *90*, 31–45. https://doi.org/10.1016/j.cortex.2017.02.004

Batterink, L. J., & Paller, K. A. (2019). Statistical learning of speech regularities can occur outside the focus of attention. *Cortex*, *115*, 56–71. https://doi.org/10.1016/j.cortex.2019.01.013

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Brodbeck, C., Hong, L. E., & Simon, J. Z. (2018). Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech. *Current Biology*, *28*(24), 3976-3983.e5. https://doi.org/10.1016/j.cub.2018.10.042

Brodbeck, C., Kandylaki, K. D., & Scharenborg, O. (2024). Neural Representations of Non-native Speech Reflect Proficiency and Interference from Native Language Knowledge. *The Journal of Neuroscience*, *44*(1), e0666232023. https://doi.org/10.1523/JNEUROSCI.0666-23.2023

Buiatti, M., Pena, M., & Dehaenelambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *NeuroImage*, *44*(2), 509–519. https://doi.org/10.1016/j.neuroimage.2008.09.015

Crosse, M. J., Di Liberto, G. M., Bednar, A., & Lalor, E. C. (2016). The Multivariate Temporal Response Function (mTRF) Toolbox: A MATLAB Toolbox for Relating Neural Signals to Continuous Stimuli. *Frontiers in Human Neuroscience*, *10*. https://doi.org/10.3389/fnhum.2016.00604

Crosse, M. J., Zuk, N. J., Di Liberto, G. M., Nidiffer, A. R., Molholm, S., & Lalor, E. C. (2021). Linear Modeling of Neurophysiological Responses to Speech and Other Continuous Stimuli: Methodological Considerations for Applied Research. *Frontiers in Neuroscience*, *15*, 705621. https://doi.org/10.3389/fnins.2021.705621

Cunillera, T., Càmara, E., Laine, M., & Rodríguez-Fornells, A. (2010). Words as Anchors: Known Words Facilitate Statistical Learning. *Experimental Psychology*, *57*(2), 134–141. https://doi.org/10.1027/1618-3169/a000017

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Desai, M., Holder, J., Villarreal, C., Clark, N., Hoang, B., & Hamilton, L. S. (2021). Generalizable EEG Encoding Models with Naturalistic Audiovisual Stimuli. *The Journal of Neuroscience*, *41*(43), 8946–8962. https://doi.org/10.1523/JNEUROSCI.2891-20.2021

Di Liberto, G. M., Nidiffer, A., Crosse, M. J., Zuk, N., Haro, S., Cantisani, G., Winchester, M. M., Igoe, A., McCrann, R., Chandra, S., Lalor, E. C., & Baruzzo, G. (2023). *A standardised open science framework for sharing and re-analysing* *neural data acquired to continuous stimuli* (Version 3). arXiv. https://doi.org/10.48550/ARXIV.2309.07671

Di Liberto, G. M., Nie, J., Yeaton, J., Khalighinejad, B., Shamma, S. A., & Mesgarani, N. (2021). Neural representation of linguistic feature hierarchy reflects second-language proficiency. *NeuroImage*, *227*, 117586. https://doi.org/10.1016/j.neuroimage.2020.117586

Elmer, S., Valizadeh, S. A., Cunillera, T., & Rodriguez-Fornells, A. (2021). Statistical learning and prosodic bootstrapping differentially affect neural synchronization during speech segmentation. *NeuroImage*, *235*, 118051. https://doi.org/10.1016/j.neuroimage.2021.118051

Erica Mandy. (n.d.). *The Newsworthy* [Broadcast].

Franco, A., Eberlen, J., Destrebecqz, A., Cleeremans, A., & Bertels, J. (2015). Rapid Serial Auditory Presentation: A New Measure of Statistical Learning in Speech Segmentation. *Experimental Psychology*, *62*(5), 346–351. https://doi.org/10.1027/1618-3169/a000295

Frank, M. C., Tenenbaum, J. B., & Gibson, E. (2013). Correction: Learning and Long-Term Retention of Large-Scale Artificial Languages. *PLoS ONE*, *8*(4). https://doi.org/10.1371/annotation/fcf19f94-5f2c-4a5f-9a6f-3d880bcdd850

Gilbert, A. C., Lee, J. G., Coulter, K., Wolpert, M. A., Kousaie, S., Gracco, V. L., Klein, D., Titone, D., Phillips, N. A., & Baum, S. R. (2021). Spoken Word Segmentation in First and Second Language: When ERP and Behavioral Measures Diverge. *Frontiers in Psychology*, *12*, 705668. https://doi.org/10.3389/fpsyg.2021.705668

Gillis, M., Vanthornhout, J., Simon, J. Z., Francart, T., & Brodbeck, C. (2021). Neural Markers of Speech Comprehension: Measuring EEG Tracking of Linguistic Speech Representations, Controlling the Speech Acoustics. *The Journal of Neuroscience*, *41*(50), 10316–10329. https://doi.org/10.1523/JNEUROSCI.0812-21.2021

Gonzalez-Gomez, N., Poltrock, S., & Nazzi, T. (2013). A "Bat" Is Easier to Learn than a "Tab": Effects of Relative Phonotactic Frequency on Infant Word Learning. *PLoS ONE*, *8*(3), e59601. https://doi.org/10.1371/journal.pone.0059601

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can Infants Map Meaning to Newly Segmented Words?: Statistical Segmentation and Word Learning. *Psychological Science*, *18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Gullberg, M., Roberts, L., Dimroth, C., Veroude, K., & Indefrey, P. (2010). Adult Language Learning After Minimal Exposure to an Unknown Natural Language. *Language Learning*, *60*(s2), 5–24. https://doi.org/10.1111/j.1467-9922.2010.00598.x

Hamilton, L. S., & Huth, A. G. (2020). The revolution will not be controlled: Natural stimuli in speech neuroscience. *Language, Cognition and Neuroscience*, *35*(5), 573–582. https://doi.org/10.1080/23273798.2018.1499946

Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, *63*(2), 93–106. https://doi.org/10.1016/j.cogpsych.2011.06.002

Honnibal, M., & Montani, I. (n.d.). *spaCy: Industrial-strength Natural Language Processing in Python* [Computer software].

Ihara, A. S., Matsumoto, A., Ojima, S., Katayama, J., Nakamura, K., Yokota, Y., Watanabe, H., & Naruse, Y. (2021). Prediction of Second Language Proficiency Based on Electroencephalographic Signals Measured While Listening to Natural Speech. *Frontiers in Human Neuroscience*, *15*, 665809. https://doi.org/10.3389/fnhum.2021.665809

Isbilen, E. S., & Christiansen, M. H. (2022). Statistical Learning of Language: A Meta-Analysis Into 25 Years of Research. *Cognitive Science*, *46*(9), e13198. https://doi.org/10.1111/cogs.13198

Kabdebon, C., Pena, M., Buiatti, M., & Dehaene-Lambertz, G. (2015). Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain and Language*, *148*, 25–36. https://doi.org/10.1016/j.bandl.2015.03.005

Karaman, F., Lany, J., & Hay, J. F. (2024). Can Infants Retain Statistically Segmented Words and Mappings Across a Delay? *Cognitive Science*, *48*(3), e13433. https://doi.org/10.1111/cogs.13433

Karunathilake, I. M. D., Kulasingham, J. P., & Simon, J. Z. (2023). Neural tracking measures of speech intelligibility: Manipulating intelligibility while keeping acoustics unchanged. *Proceedings of the National Academy of Sciences*, *120*(49), e2309166120. https://doi.org/10.1073/pnas.2309166120

Kisler, T., Reichel, U., & Schiel, F. (2017). Multilingual processing of speech via web services. *Computer Speech & Language*, *45*, 326–347. https://doi.org/10.1016/j.csl.2017.01.005

Kittleson, M. M., Aguilar, J. M., Tokerud, G. L., Plante, E., & Asbjørnsen, A. E. (2010). Implicit language learning: Adults' ability to segment words in Norwegian. *Bilingualism: Language and Cognition*, *13*(4), 513–523. https://doi.org/10.1017/S1366728910000039

Lalor, E. C., & Foxe, J. J. (2010). Neural responses to uninterrupted natural speech can be extracted with precise temporal resolution. *European Journal of Neuroscience*, *31*(1), 189–193. https://doi.org/10.1111/j.1460-9568.2009.07055.x

Lany, J., Karaman, F., & Hay, J. F. (2024). A changing role for transitional probabilities in word learning during the transition to toddlerhood? *Developmental Psychology*. https://doi.org/10.1037/dev0001641

Liu, H., Forest, T. A., Duncan, K., & Finn, A. S. (2023). What sticks after statistical learning: The persistence of implicit versus explicit memory traces. *Cognition*, *236*, 105439. https://doi.org/10.1016/j.cognition.2023.105439

Lukics, K. S., & Lukács, Á. (2021). Tracking statistical learning online: Word segmentation in a target detection task. *Acta Psychologica*, *215*, 103271. https://doi.org/10.1016/j.actpsy.2021.103271

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, *59*(4), 475–494. https://doi.org/10.1016/j.jml.2007.11.006

Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition*, *108*(1), 271–280. https://doi.org/10.1016/j.cognition.2008.02.003

*News In Slow Italian*. (n.d.). [Broadcast].

Oh, Y., Todd, S., Beckner, C., Hay, J., King, J., & Needle, J. (2020). Non-Māori-speaking New Zealanders have a Māori proto-lexicon. *Scientific Reports*, *10*(1), 22318. https://doi.org/10.1038/s41598-020-78810-4

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009). Statistical Learning in a Natural Language by 8-Month-Old Infants. *Child Development*, *80*(3), 674–685. https://doi.org/10.1111/j.1467-8624.2009.01290.x

Pérez-Navarro, J., Klimovich-Gray, A., Lizarazu, M., Piazza, G., Molinaro, N., & Lallier, M. (2024). Early language experience modulates the tradeoff between acoustic-temporal and lexico-semantic cortical tracking of speech. *iScience*, *27*(7), 110247. https://doi.org/10.1016/j.isci.2024.110247

Plante, E., & Gómez, R. L. (2018). Learning Without Trying: The Clinical Relevance of Statistical Learning. *Language, Speech, and Hearing Services in Schools*, *49*(3S), 710–722. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0131

Poulin-Charronnat, B., Perruchet, P., Tillmann, B., & Peereman, R. (2017). Familiar units prevail over statistical cues in word segmentation. *Psychological Research*, *81*(5), 990–1003. https://doi.org/10.1007/s00426-016-0793-y

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *WIREs Cognitive Science*, *1*(6), 906–914. https://doi.org/10.1002/wcs.78

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical Learning by 8-Month-Old Infants. *Science*, *274*(5294), 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word Segmentation: The Role of Distributional Cues. *Journal of Memory and Language*, *35*(4), 606–621. https://doi.org/10.1006/jmla.1996.0032

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental Language Learning: Listening (and Learning) Out of the Corner of Your Ear. *Psychological Science*, *8*(2), 101–105. https://doi.org/10.1111/j.1467-9280.1997.tb00690.x

Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting Cues: Overlapping Regularities Support Cue Discovery in Infancy. *Child Development*, *81*(3), 727–736. https://doi.org/10.1111/j.1467-8624.2010.01430.x

Sanders, L. D., & Neville, H. J. (2003a). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, *15*(3), 228–240. https://doi.org/10.1016/S0926-6410(02)00195-7

Sanders, L. D., & Neville, H. J. (2003b). An ERP study of continuous speech processing: II. Segmentation, semantics, and syntax in non-native speakers. *Cognitive Brain Research*, *15*(3), 214–227. https://doi.org/10.1016/S0926-6410(02)00194-5

Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: An event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, *5*(7), 700–703. https://doi.org/10.1038/nn873

Sassenhagen, J. (2019). How to analyse electrophysiological responses to naturalistic language with time-resolved multiple regression. *Language, Cognition and Neuroscience*, *34*(4), 474–490. https://doi.org/10.1080/23273798.2018.1502458

Shoaib, A., Wang, T., Hay, J. F., & Lany, J. (2018). Do Infants Learn Words From Statistics? Evidence From English-Learning Infants Hearing Italian. *Cognitive Science*, *42*(8), 3083–3099. https://doi.org/10.1111/cogs.12673

Sjuls, G. S., Harvei, N. N., & Vulchanova, M. D. (2023). The relationship between neural phase entrainment and statistical word-learning: A scoping review. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-023-02425-9

Sohail, J., & Johnson, E. K. (2016). How Transitional Probabilities and the Edge Effect Contribute to Listeners' Phonological Bootstrapping Success. *Language Learning and Development*, *12*(2), 105–115. https://doi.org/10.1080/15475441.2015.1073153

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*(6640), 381–382. https://doi.org/10.1038/41102

The Audacity Team. (n.d.). *Audacity(R)* (Version 3.3.2.0) [Computer software].

TIME. (n.d.). *TIME's The Brief* [Broadcast].

Urbain, C., Bourguignon, M., Op De Beeck, M., Schmitz, R., Galer, S., Wens, V., Marty, B., De Tiège, X., Van Bogaert, P., & Peigneux, P. (2013). MEG Correlates of

Learning Novel Objects Properties in Children. *PLoS ONE*, *8*(7), e69696. https://doi.org/10.1371/journal.pone.0069696

Vanthornhout, J., Decruy, L., & Francart, T. (2019). Effect of Task and Attention on Neural Tracking of Speech. *Frontiers in Neuroscience*, *13*, 977. https://doi.org/10.3389/fnins.2019.00977

Wojcik, E. H., Zettersten, M., & Benitez, V. L. (2022). The map trap: Why and how word learning research should move beyond mapping. *WIREs Cognitive Science*, *13*(4), e1596. https://doi.org/10.1002/wcs.1596

Zou, J., Feng, J., Xu, T., Jin, P., Luo, C., Zhang, J., Pan, X., Chen, F., Zheng, J., & Ding, N. (2019). Auditory and language contributions to neural encoding of speech features in noisy environments. *NeuroImage*, *192*, 66–75. https://doi.org/10.1016/j.neuroimage.2019.02.047

# Appendix A: Excerpt from Rae Hoeppner's Thesis

***Model Training and Calculating Phonotactic Score.***

To calculate the phonotactic probability of each word, the exposure transcripts were converted into their phonetic components in IPA and concatenated into a single dataset. Punctuation, including hyphens, was removed to treat the transcript as a continuous, unsegmented stream, except for paragraph breaks denoting speaker or podcast changes. Phonotactic models were trained on the Italian exposure podcasts using the SRI Language Modeling Toolkit (SRILM), a specialized toolkit for statistical language models (Stolcke, 2002). A trigram model with Whitten-Bell smoothing and no tokenization was trained on the exposure podcast dataset, containing the sequence of phonemes of all the exposure podcasts. This was used to then generate log probabilities of the phonetic sequence of each test item using the trained model. The resulting log probabilities of the phonetic sequence of each word were then normalized by dividing them by the number of phonemes in the sequence plus one, accounting for the end-of-word symbol (Oh et al., 2020). This normalization process yielded the phonotactic score of each word. A higher (less negative) score signifies a higher likelihood of the phonetic sequence occurring in Italian. Finally, our two frequency conditions were segregated into low and high phonotactic score conditions based on median split by phonotactic score. This resulted in a 2x2 design for items in the word category, corresponding to high and low frequency by high and low phonotactic scores in which each cell held a total of 60 items (Table 1) for a total of 240 words.

***Nonword Creation.***

Non-words were generated following the non-tokenization method outlined by Oh et al. (2020). Nonwords were found neither within the exposure podcast, nor the actual Italian language. This approach involved selecting words from our real word condition and changing up to 3 phonemes within the sequence. All nonwords had a corresponding word that was chosen at random. These alterations were made to produce words with lower phonotactic scores than all real words, resulting in phonetic sequences with an extremely low probability of occurring in the Italian language, particularly in bisyllabic words.

### ***Text-to-Speech Procedure.***

The phonetic transcripts of each word and non-word were converted into orthographic transcripts using the BAS Web Services Pho2Syl service (Reichel, 2012). This service translates phonological transcripts into syllables, which were then concatenated to create the orthographic representation of each word. Subsequently, these orthographic representations were adjusted to ensure correct pronunciation by the text-to-speech program. Audio files for each word were then generated using the Google Cloud Console's text-to-speech functionality. The audio files were saved in WAV format at a sampling rate of 48000 Hz. Finally, a phoneme-to-speech website, the IPA Reader with the setting "Carla [Italian]" (Linero, 2018) which uses Amazon's Polly text-to-speech service, was utilized to verify the accuracy of the Cloud Console's generated words. The Google Cloud Console words were auditorily checked against the IPA reader for all test items and the orthography of nonwords were changed to fit the phonetic representation. This step was necessary because Google Cloud Console software requires orthographic input, and thus, the phonetic transcripts needed to be converted into their corresponding orthographic representations. However, the generated orthographic input did not always suitably represent the nonword, or the software could not suitably pronounce the nonword, and therefore the orthography was changed to reflect the phonetic pronunciation.

To ensure that the created non-words did not inadvertently resemble real Italian words, all nonword audio files were reviewed by a native Italian speaker. Within this step, the native Italian speaker confirmed that nonwords resembled natural Italian. Any nonword that was accidently a real Italian word, or sounded similar to real Italian words was flagged and subsequently replaced with another nonword. This step ensured the integrity of the nonword stimuli.

# Appendix B: Word Mapping Task Stimuli

| ***Version A*** | ***Word Type*** | ***Phonotactic Score*** | ***Frequency*** |
|---|---|---|---|
| zanne | rare | -1.72881 | 1 |
| ceppo | rare | -1.7223 | 1 |
| stuoie | rare | -1.79374 | 1 |
| doccia | rare | -1.33171 | 1 |
| anno | common | -0.61672 | 81 |
| festa | common | -0.78779 | 19 |
| centro | common | -0.78647 | 71 |
| maggio | common | -0.74985 | 40 |
| chispo | non-word | -3.14435 | 0 |
| gliona | non-word | -2.43645 | 0 |
| djagna | non-word | -2.39455 | 0 |
| irlo | non-word | -2.23535 | 0 |
| | | | |
| ***Version B*** | ***Word Type*** | ***Phonotactic Score*** | ***Frequency*** |
| ghiaccio | rare | -1.45243 | 1 |
| mappe | rare | -1.26506 | 1 |
| righe | rare | -1.40363 | 1 |
| targhe | rare | -1.27331 | 1 |
| mare | common | -0.70774 | 32 |
| posto | common | -0.72419 | 46 |
| tratti | common | -0.79291 | 7 |
| volta | common | -0.6984 | 91 |
| camlo | non-word | -2.36862 | 0 |
| gnevu | non-word | -4.18534 | 0 |
| srellio | non-word | -2.4507 | 0 |
| tjutse | non-word | -2.86532 | 0 |

# Appendix C: Self-reported participant activities during listening period

Total greater than 100% because some participants reported multiple activities.

| **Self-Reported Activity** | **Percentage of Participants** |
|---|---|
| Cleaning | 76.74% |
| Cooking/Eating | 65.12% |
| Commuting | 76.74% |
| Consuming Media | 9.30% |
| Getting Ready/Showering | 9.30% |
| Exercising | 6.98% |
| Art | 4.65% |

# Appendix D: Ethics Approval

**Date:** 29 March 2023

**To:** Dr Laura Batterink

**Project ID:** 122155

**Study Title:** Effect of passive exposure to a second language on speech processing

**Short Title:** Passive exposure to L2

**Application Type:** NMREB Initial Application

**Review Type:** Delegated

**Full Board Reporting Date:** 14/Apr/2023

**Date Approval Issued:** 29/Mar/2023 16:51

**REB Approval Expiry Date:** 29/Mar/2024

---

Dear Dr Laura Batterink

The Western University Non-Medical Research Ethics Board (NMREB) has reviewed and approved the WREM application form for the above mentioned study, as of the date noted above. NMREB approval for this study remains valid until the expiry date noted above, conditional to timely submission and acceptance of NMREB Continuing Ethics Review.

This research study is to be conducted by the investigator noted above. **All other required institutional approvals and mandated training must also be obtained prior to the conduct of the study.**

**Documents Approved:**

| Document Name | Document Type | Document Date | Document Version |
|---|---|---|---|
| Debriefing_L2 | Debriefing document | 22/Dec/2022 | 1 |
| English_podcast_1 | Other Data Collection Instruments | 22/Dec/2022 | 1 |
| Italian_podcast_1 | Other Data Collection Instruments | 22/Dec/2022 | 1 |
| Participant Information L2 | Paper Survey | 22/Dec/2022 | 1 |
| Word Mapping L2 | Other Data Collection Instruments | 22/Dec/2022 | 1 |
| Word Rating L2 | Other Data Collection Instruments | 22/Dec/2022 | 1 |
| Daily_podcast_questionnaire | Online Survey | 24/Feb/2023 | 2 |
| Flyer_L2 | Recruitment Materials | 21/Mar/2023 | 3 |
| Email_script_L2 | Recruitment Materials | 21/Mar/2023 | 3 |
| BrainsCAN_L2 | Recruitment Materials | 21/Mar/2023 | 3 |
| Consent_L2 | Written Consent/Assent | 21/Mar/2023 | 3 |

**Documents Acknowledged:**

| Document Name | Document Type | Document Date | Document Version |
|---|---|---|---|
| Eligibility_Checklist_L2 | Screening Form/Questionnaire | 23/Feb/2023 | 2 |

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

Sincerely,

Ms. Katelyn Harris , Research Ethics Officer on behalf of Dr. Randal Graham, NMREB Chair

Date: 19 March 2024

**To:** Dr Laura Batterink

**Project ID:** 122155

**Study Title:** Effect of passive exposure to a second language on speech processing

**Application Type:** Continuing Ethics Review (CER) Form

**Review Type:** Delegated

**Date Approval Issued:** 19/Mar/2024 15:22

**REB Approval Expiry Date:** 29/Mar/2025

---

Dear Dr Laura Batterink,

The Western University Non-Medical Research Ethics Board has reviewed this application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

The Western University NMREB operates in compliance with the Tri-Council Policy Statement Ethical Conduct for Research Involving Humans (TCPS2), the Ontario Personal Health Information Protection Act (PHIPA, 2004), and the applicable laws and regulations of Ontario. Members of the NMREB who are named as Investigators in research studies do not participate in discussions related to, nor vote on such studies when they are presented to the REB. The NMREB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000941.

Please do not hesitate to contact us if you have any questions.

**Electronically signed by:**

Mr. Joshua Hatherley, Ethics Coordinator on behalf of Dr. Isha DeCoito, NMREB Chair 19/Mar/2024 15:22

**Reason:** I am approving this document

***Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).***

# Curriculum Vitae

| | |
|---|---|
| **Name:** | Amiya Aggarwal |
| **Post-secondary Education and Degrees:** | University of Waterloo<br>Waterloo, Ontario, Canada<br>2013-2017 B.Sc. Biomedical Science (Hons) |
| | University of Waterloo<br>Waterloo, Ontario, Canada<br>2018-2021 B.A. Psychology (Hons) |
| | The University of Western Ontario<br>London, Ontario, Canada<br>2022-2024 M.Sc. Psychology |
| **Honours and Awards:** | Ontario Graduate Scholarship (OGS)<br>2022-2023 |
| **Related Work Experience** | Teaching Assistant<br>The University of Western Ontario<br>2022-2024 |

**Conference Poster Presentations:**

Aggarwal. A. S., Walker, A. C., & Fugelsang. J. A. (June 2021). *Source Information and Doublespeak: How Source Cues Influence Moral Judgments.* Annual Ontario Psychology Undergraduate Thesis Conference, Canada

Aggarwal, A. S., Hoeppner, É. R., Batterink, L. J. (March 2023). *Can statistical learning support speech segmentation of a natural language in adult learners?* Cognitive Neuroscience Society Conference, San Francisco, CA, USA

Hoeppner, É. R., Aggarwal, A. S., & Batterink, L. J. (March 2023). *Harnessing Statistical Learning to Support the Discovery of Second Language Phonetic Patterns in Adult Learners*. Cognitive Neuroscience Society Conference, San Francisco, CA, USA

Aggarwal, A. S., Hoeppner, É. R., & Batterink, L. J. (February 2024). *Investigating the Impact of Passive Second Language Exposure on Speech Segmentation and Word Learning.* Lake Ontario Visionary Establishment Conference, Niagara Falls, ON, Canada

Hoeppner, É. R., Aggarwal, A. S., Batterink, L. J. (February 2024). *Harnessing Statistical Learning for Second Language Acquisition*. Lake Ontario Visionary Establishment Conference, Niagara Falls, ON, Canada

Hoeppner, É. R., Aggarwal, A. S., Batterink, L. J. (April 2024). *Harnessing Statistical Learning for Second Language Acquisition*. Cognitive Neuroscience Society Conference, Toronto, ON, Canada

Aggarwal, A. S., Hoeppner, É. R., & Batterink, L. J. (April 2024). *Impact of Passive Second Language Exposure on Speech Segmentation and Word Learning*. Cognitive Neuroscience Society Conference, Toronto, ON, Canada