

**Effects of Auditory Processing, Memory, and Experience on Early and Later Stages of  
Second Language Speech Learning****Linda Bakkouche <sup>a</sup>, \*Kazuya Saito <sup>b</sup>**<sup>a</sup> *University of Cambridge*<sup>b</sup> *University College London*\* *Corresponding Author*

**Acknowledgement**

This work originates from the first author's (LB) master's dissertation, submitted to University College London in 2021. We extend our gratitude to all participants for their involvement and to the SLR reviewers and associate editor, Guilherme Garcia, for their constructive feedback. The development and validation of the test materials were supported by the Leverhulme Trust Research Grant (RPG-2024-391) and the UK-ISPF Research Grant (1185702223), both awarded to the second author (KS).

**Abstract**

The current study aimed to explore how individual differences in lower-order perceptual abilities (auditory processing) and higher-order cognitive abilities (declarative and procedural memory) could be linked to various dimensions of L2 speech production (segmentals, prosody, fluency, and lexicogrammar) in a group of forty-seven late Spanish-English bilinguals with varying levels of L2 proficiency and immersion experience. The statistical analyses revealed several key findings: (a) auditory processing showed significant correlations with L2 speech proficiency, particularly among learners with short-to-mid immersion experience (Length Of Residence [LOR] =1-5 years); (b) a robust triangular relationship ( $r = .7-.8$ ) was observed between procedural memory, the duration of immersion experience, and L2 speech outcomes across the entire participant group (LOR= 1- 20 years); and (c) declarative memory did not exhibit any significant associations with the factors related to experience and proficiency. Based on these results, it is tentatively suggested that auditory processing abilities may support the development of L2 speech in the initial stages of immersion (LOR=1-5 years), while procedural memory skills may play a crucial role in achieving high-level L2 speech proficiency, as individuals have more opportunities to practice L2 pronunciation through active usage during an extended stay in a country where the L2 is spoken.

*Keywords:* aptitude, second language acquisition, second language speech, auditory processing, declarative memory, procedural memory

## Introduction

Over the past five decades, scholars have been actively involved in an extensive discourse examining factors such as the age of second language (L2) acquisition and experience (Munro, Derwing, & Saito, 2024) that influence L2 acquisition after puberty, further highlighting the crucial role of aptitude in attaining advanced levels of L2 proficiency (Skehan, 2016).

Traditionally, aptitude was operationalized as higher-order cognitive abilities, such as working memory and attentional control (for a comprehensive overview of aptitude, see Li, 2016). While earlier research has primarily focused on the impact of such cognitive abilities and declarative and procedural memory on the morphosyntax aspects of L2 learning, as evidenced in studies by Hamrick (2015) and Morgan-Short et al. (2014), the specific contribution of these abilities to mastering L2 speech remains unclear.

In recent years, there has been a noticeable shift towards investigating the significance of lower-order perceptual and sensory abilities (i.e., auditory processing) in L2 speech development (Kachlicka, Saito, & Tierney, 2019; Saito, Sun, & Tierney, 2020b). While some argue that auditory processing can be considered as a dynamic interplay of perceptual, cognitive, and motor abilities (Kraus & Banai, 2007; Saito et al., 2024), this ability has been essentially operationalized as a bottom-up processing of basic acoustic information (pitch, duration, and intensity).

There is emerging evidence that both perceptual and cognitive individual differences (auditory processing and memory) interact to influence the outcomes of L2 speech perception (Saito et al., 2022a). We have yet to know the complex relationship between auditory processing, memory and L2 speech production. The following study, involving 47 Spanish-English bilinguals with diverse experience and proficiency levels, aims to unravel the intricate relationship between experience, auditory processing, and memory in the multifaceted dimensions of L2 speech proficiency, with a specific emphasis on phonology and lexicogrammar.

## Background

### Auditory Processing

In the context of first language (L1) acquisition, Kraus and Banai (2007) proposed the interaction view that auditory processing is a multifaceted phenomenon, a complex interplay of perceptual, cognitive, and motor abilities operating on different levels (for the context of L2 acquisition, see Saito et al., 2024). In general, however, auditory processing has been considered as a relatively low-order, perceptual ability to encode acoustic details of sounds.

Temporal acuity refers to an individual's ability to discern changes in amplitude over time. This ability is crucial for diverse aspects related to fluency, such as phonation time and pause frequency, and segmental contrasts like short versus long vowels. Temporal acuity also impacts prosody, including durational differences in weak versus strong vowels, voice onset time, and rhythm, such as the duration between two syllables or stressed syllables versus morae (Tallal et al, 1993). Spectral processing, on the other hand, pertains to an individual's ability to track changes in the frequency content of a signal, such as pitch (the frequency of vocal cord vibration) and formants (the energy concentration at different frequency bands). This ability is essential for the appropriate assignment of stress and intonation, fine-tuning to isochrony (e.g., syllable-timed vs. stressed-timed), and the refinement of segmental accuracy (e.g., third formant variability for English [r] and [l]) (Saito et al., 2022b).

Auditory processing is a pivotal element in the acquisition of an L1, and variations in auditory processing capabilities can significantly influence the pace at which L1 is acquired. As children evolve through the process of L1 acquisition, they become increasingly proficient in distinguishing temporal and spectral cues within their native language(s). However, these improvements do not necessarily translate to the perception of non-native language features,

with a plateau often being reached by the end of their first year. Auditory acuity, a key aspect of auditory processing, develops and typically peaks around the ages of 7 to 10. Following this peak, there is a gradual decline in the precision of auditory acuity as individuals age (Skoe, Krizman, Anderson, & Kraus, 2015). Impediments in auditory processing have been linked to a variety of deficits. These include language skills, attention, memory, and reading difficulties, as observed in conditions like dyslexia (Surprenant & Watson, 2001). Moreover, differences in auditory processing at an individual level have been associated with challenges in L1 learning (Goswami et al., 2011). Despite the substantial body of evidence supporting these connections, the causal relationship between auditory processing and L1 acquisition remains a subject of ongoing debate (cf., Rosen & Manganari, 2001; Snowling et al., 2018).

Auditory processing is considered fundamental in L2 speech learning (Kachlicka et al., 2019, Saito et al., 2024) as presented in *Table 1*. Research, including studies by Kempe et al. (2012), Chandrasekaran et al. (2010), Perrachione et al. (2011), and Wong & Perrachione (2007), supports the notion that individuals with higher perceptual acuity are better at perceiving and learning unfamiliar sounds. This predictive power of acuity is particularly pronounced in naturalistic L2 learners who have achieved proficiency in a target language through immersive experiences. Specifically, perceptual acuity is a moderate to strong predictor of successful L2 speech learning outcomes for those with medium to long-term immersion experiences (residing for 1-10 years; Kachlicka et al., 2019; Saito et al., 2020a). Recent findings by Saito et al. (2024) underscore the significant role of auditory processing in predicting L2 perception accuracy in both phonology and lexicogrammar. However, it is less predictive for those with short-term immersion experiences (less than 6 months; Saito et al., 2022c) or for foreign language learners with no immersion experience (Saito et al., 2021).

Our research aims to further investigate these relationships, with a specific focus on auditory acuity, particularly temporal and spectral processing, which, as suggested by previous research, may play a critical role in L2 speech learning. In doing so, we aim to elucidate the intricate interplay between auditory processing and the mechanisms of other cognitive abilities (i.e., procedural and declarative memory) in the context of L2 acquisition. By understanding these connections, we can also provide valuable insights into the cognitive processes underlying successful L2 speech learning.

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**Table 1***Summary of Key Studies in Auditory Processing and Second Language Acquisition*

Studies	Subjects	Task(s)	Measures	Findings
Legeris & Hazan (2010)	28 Greek speakers of English	Perception task- non-native (English) vowel discrimination task	Auditory acuity (spectral - formant discrimination)	Individuals with more accurate auditory processing demonstrated more accurate L2 speech after they had received the speech training
Kachlicka et al. (2019)	40 Polish speakers of English (LOR = 1-5 years)	Perception task – non-native (English) vowel discrimination Grammaticality judgment task	Auditory acuity (spectral and temporal processing)  Integration (reproduction)	Spectral processing skills were linked to more precise L2 speech perception and grammatical judgement.
Saito et al. (2020b) a longitudinal study	30 Chinese speakers of English (LOR = 8 months)	Production- picture description task (segmental, prosody and fluency assessment)	Auditory acuity (spectral and temporal assessment)	Individuals with more precise auditory processing skills showed more accurate segmental and suprasegmental dimensions of their L2 speech improving with the length of residence.

(continued)

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**Table 1** (continued)*Summary of Key Studies in Auditory Processing and Second Language Acquisition*

Studies	Subjects	Task(s)	Measures	Findings
Saito et al. (2020a)	139 L2 speakers: Chinese ( $n = 50$ , LOR = 4 months) Polish ( $n = 50$ , LOR = 6-19 years) Spanish ( $n = 39$ , LOR = 2-24 years)	Perception task: vowel and prosody discrimination Grammaticality judgement task	Auditory acuity (spectral and temporal assessment)	L2 learners with longer immersion experience (LOR > 1 year) and better auditory processing skills demonstrated more fluent and accurate L2 speech.
Zheng et al. (2020)	48 Chinese speakers of English (LOR = 1 year)	Production- picture description task	Auditory acuity (spectral and temporal assessment)	Auditory processing (both spectral and temporal processing) abilities predicted a high degree of attainment in precise segmentals and prosody
Sun et al. (2021)	46 Chinese speakers of English (LOR = 5 months)	Perception task- non-native (English) vowel and consonant discrimination	Auditory acuity (spectral and temporal assessment)- explicit auditory processing measures Electrophysiological measures - implicit auditory processing measures	Individuals with greater auditory acuity (explicit measure) presented better non-native sound discrimination

*Note.* LOR = Length of Residence.

### **Declarative and Procedural Memory**

Ullman's (2004) Declarative/Procedural (D/P) model emphasizes the essential role of these long-term memory systems in language learning, storage, and utilization, serving as crucial domain-general learning mechanisms (Hamrick, Lum, & Ullman, 2018). Notably, these memory systems differ in various dimensions, including their relationship with awareness, their role in language learning, and the associated brain regions (Eichenbaum, 2002).

Declarative memory, linked to the hippocampus and its surrounding structures (Poldrack & Packard, 2003), is primarily responsible for the explicit learning of episodic and semantic knowledge and associations (Tulving, 1993). This memory system supports the acquisition of events (episodic memory), facts (semantic memory), and the storage of lexical items crucial for all idiosyncratic linguistic knowledge at the word or multi-word level. In contrast, procedural memory, crucial for implicit learning, encompasses a range of perceptual-motor and cognitive skills like navigation (Ullman & Lovelett, 2018). Despite its slower learning pace and need for repeated exposure, knowledge acquired in procedural memory is processed more rapidly and automatically than in declarative memory (Ullman, 2020). Moreover, procedural memory tends to be more robust in early childhood and diminishes in early adulthood, as evidenced by Nameth, Janacek, and Fiser (2013).

The interaction between declarative and procedural memory systems in L2 learning is a complex and dynamic process as presented in *Table 2*. Studies have shown that these memory systems influence L2 learners' linguistic outcomes differently based on their stage of acquisition, with declarative memory abilities predicting early-stage L2 grammar learning and individuals with stronger procedural learning abilities showing larger gains with increased L2 input (Hamrick, 2015; Morgan-Short et al., 2014). Research in classroom settings has explored how these memory systems relate to L2 speech learning, with findings suggesting that greater procedural learning ability is associated with L2 speech fluency (Granena, 2019). Recent studies, such as that by Quam et al. (2018) demonstrated that those with greater declarative memory benefited more from short-term training, whilst those with greater procedural memory yielded more gains as a result of an increased length of training. More recently, Saito (2022a) further highlight the importance of procedural memory for precise L2 speech perception among Japanese learners of English in foreign language settings.

However, the extent to which long-term memory abilities can support L2 learners in immersion environments remains to be fully understood. Given the intricacies of these memory systems and their roles in L2 learning, additional research is warranted to enhance our understanding of how they interact, particularly among experienced and proficient L2 speakers in naturalistic settings.



## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**Table 2***Summary of Key Studies in Declarative and Procedural Memory and Second Language Acquisition*

Studies	Subjects	Task(s)	Measures	Findings
Linck et al. (2013)	522 L1 speakers of English	Reading and Listening Task in L2	Declarative memory- Paired Associates Implicit (procedural memory learning) – SRT task	The greater explicit (declarative) and implicit learning ability predicted listening and reading attainment.
Morgan-Short et al. (2014) a laboratory study	14 L1 speakers of English	Grammaticality judgement task (artificial language)	Declarative memory: The Modern Language Aptitude Test, Part V, The Continuous Visual Memory Task  Procedural memory: The Weather Prediction Task, The Tower of London Task	Declarative memory ability predicted L2 syntactic structure development at early stages of acquisition.  Procedural memory ability predicted L2 syntactic structure development at later stages of acquisition.
Hamrick (2015) a laboratory study	31 L1 speakers of English	The semi-artificial language L2 syntactic structure learning	Declarative memory - LLAMA B Procedural memory - SRT Task	Learners with better declarative memory learning showed greater gains in the immediate recognition task. Learners with better procedural memory learning demonstrated greater gains in the delayed recognition task.

(continued)

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**Table 2** (continued)*Summary of Key Studies in Declarative and Procedural Memory and Second Language Acquisition*

Studies	Subjects	Task(s)	Measures	Findings
Brooks et al. (2017) a laboratory study	44 learners of L2 - Russian	Auditory artificial grammar learning task	Procedural memory - Artificial grammar learning task	Statistical learning ability and non-verbal intelligence were significant predictors of L2 morphology accuracy.
Quam et al. (2018) a laboratory study	29 adult native speakers of English	Novel sound (vowels) category learning	Procedural memory was assessed with the SRT task Declarative memory was assessed with the subtest of the Wechsler Memory Scale- 4th edition	Declarative memory ability predicted accuracy and cue integration at the initial (but not later) stages of learning. Procedural memory ability predicted accuracy and cue integration in an auditory category
Granena (2019) Foreign Language (FL) Setting	135 L1 speakers of English, learning L2 - Spanish	L2 production -the picture description task	Explicit (declarative) learning: LLAMA_B Implicit (procedural) learning: SRT Task	Learners with greater procedural memory learning ability demonstrated higher L2 speech fluency.
Saito et al. (2021) FL Setting	70 Japanese learners of L2-English	Two-alternative forced-choice identification test- phonetic contrasts	Declarative memory assessed by LLAMA_B Procedural Memory- SRT task	Procedural memory predicted L2 speech perception accuracy

*Note.* SRT = Serial Reaction Time.

## **L2 Speech Learning**

L2 speech proficiency has been conceptualized as a composite construct consisting of various linguistic domains such as the correct pronunciation of individual sounds (segmentals), the adequate and varied use of stress (prosody), and the accurate use of lexes and morphosyntactic markers (lexicogrammar; e.g., see Saito & Plonsky, 2019, for a methodological review on the outcome measures of L2 speech proficiency). There is ample cross-sectional and longitudinal research showing how the outcomes of adult L2 speech learning can be associated with the quantity and quality of experience at two different stages, i.e., rate of learning vs. ultimate attainment (see DeKeyser, 2013, for a comprehensive review).

During the initial years of immersion, L2 learners often experience significant improvement in various aspects of L2 speech as they actively seek and engage in conversations. This stage of L2 learning is referred to as the “rate of learning,” where individuals enhance their proficiency based on factors such as the length, timing, and intensity of immersion experiences (Munro, Derwing, & Saito, 2013 for phonology; Saito, 2015 for lexicogrammar; Segalowitz & Freed, 2004 for fluency).

After extensive immersion experience, L2 learners’ proficiency tends to reach a more stable or plateaued state, becoming less susceptible to further practice opportunities. This stage, known as “ultimate attainment,” suggests that experience-related factors alone cannot fully explain the outcomes of L2 speech learning. Even among experienced L2 learners, there is considerable individual variation, with some achieving native-like proficiency and others maintaining a noticeable foreign accent (Derwing & Munro, 2013). Further research is needed to explore how learner-internal, aptitude-related factors contribute to the development of high-level L2 oral proficiency (Granena, 2013).

### **Motivation for the Current Study**

To date, a growing amount of evidence has pointed out the broad relationship between auditory processing, long-term memory systems (i.e., declarative and procedural), experience, and L2 speech proficiency (Granena, 2019; Kachlicka et al., 2019; Linck et al., 2013; Saito et al., 2020a). However, it remains unclear how these variables interact with each other and to what extent they support adult learners in producing more accurate speech in immersive environments. First, few studies have probed the relative weights of auditory processing, long-term memory, and learners’ experience-related factors in L2 speech learning (cf. Saito et al., 2022a).

Second, Skehan’s (2016) aptitude-acquisition model has argued (a) that auditory processing could be a key factor during the early phase of L2 acquisition as it determines the amount of input that learners can take in; and (b) that declarative and procedural memory may relate to the incidence of high-level L2 acquisition as the long-term memory abilities predict the extent to which learners can store, transform, and elaborate received input with a view of sophisticated linguistic analyses. Therefore, it is crucial to determine the potential varying strengths of the relationship between auditory processing, the two long-term memory systems, and L2 speech outcomes among L2 learners across different immersion and proficiency levels. By doing so, we can gain a more comprehensive understanding of these complex interrelationships and their impact on L2 learning. This exploration is vital as it will provide valuable insights into how these factors interact and contribute to the progress and success of L2 learners.

Finally, while the existing literature has shown that auditory processing (spectral and temporal acuity), long-term (declarative and procedural) memory and L2 experience (i.e., Length of Residence [LOR], Age on Arrival [AOA], Current L2 Use and Foreign Language Education) are predictive of L2 speech learning respectively, very few studies have expounded the inter-relationships between the eight predictor variables in the context of post-

pubertal L2 users. There is some evidence that both auditory processing and procedural memory are susceptible to chronological age (older participants likely have less auditory precision and memory; see Nameth et al., 2013, for procedural memory; Saito et al., 2020a, for auditory processing). Little is known about how perceptual-cognitive abilities can be influenced by the length, timing, and intensity of the immersion experience.

The current study seeks to examine the extent to which aptitude and experience factors contribute to more accurate production across multiple dimensions of L2 speech, according to the framework developed and proposed by Saito and Plonsky (2019). These dimensions include vowels, consonants, word stress, intonation, rhythm, speech rate, semantic accuracy, and morphosyntactic accuracy. By examining how they interact and their contribution to variances in the segmental, prosody, fluency, and lexicogrammar aspects of L2 speech proficiency, we hope to provide a more in-depth understanding of L2 speech learning. This exploration is particularly relevant for Spanish speakers of English with varied levels of proficiency and immersion experience in the UK (1-20 years). The understanding derived from this study could inform tailored teaching strategies and methodologies that take into account these key variables. Therefore, the following research questions were formulated to guide this investigation:

1. How do auditory processing, procedural and declarative memory, and factors related to the experience of L2 learners (Length of Residence [LOR], Age on Arrival [AOA], Current L2 Use, Foreign Language Education) interact with each other?
2. To what extent do experience factors and individual differences in auditory processing and two memory systems predict the attainment of four dimensions of L2 speech: segmentals (vowels and consonants), prosody (word stress and intonation), fluency (rhythm and speech rate), and lexicogrammar (lexical and morphosyntactic accuracy)?

As for RQ1, both auditory processing (Saito et al., 2020b; Skoe, Krizman, Anderson, & Kraus, 2013) and declarative memory (Ullman & Lovelett, 2018) tend to decline with age. Given this evidence, we anticipated finding a similar relationship in the current study. Furthermore, based on existing research, we hypothesized that auditory processing would not be influenced by the extent of language experience (Saito et al., 2024). This stands in contrast to procedural memory aptitude, which previous studies have found to be related to the length of residence in a L2 speaking country (Saito et al., 2021). Consequently, our study aims to add to this body of knowledge by exploring these relationships further.

For RQ 2, we predicted that both auditory processing and long-term memory system (more procedural than declarative memory) could play equally significant roles in L2 speech learning. We postulated that learners with higher auditory acuity and memory capabilities might exhibit more precise prosody, segmentals, and lexicogrammar, as suggested by Granena (2019) and Saito et al. (2021).

Considering that declarative or explicit memory learning is believed to be associated with the initial stages of adult learning (Ullman, 2020), we anticipated that our study participants who have varying levels of immersion experience (LOR = 1-20 years) would demonstrate a greater reliance on procedural memory. This cognitive ability, which is thought to be characteristic of the later stages of L2 learning, is often enhanced with language experience (Hamrick, 2015; Morgan-Short et al., 2014; Suzuki & DeKeyser, 2015; Ullman, 2020). Therefore, we hypothesized that participants with more extensive L2 immersion experience would display a higher dependence on procedural memory, reflecting its role in facilitating more advanced stages of language acquisition.

### **Method**

All the research materials used in the current project have been deposited on the open science platform for researchers and teachers, L2 Speech Tools (Mora-Plaza et al., 2021).

## Participants

The participants in the current study were 47 Spanish speakers of English who were studying or working in the UK. Since the Catalan language is spoken in certain regions of Spain (e.g., in Catalonia) in addition to the official language of Spanish, Spanish bilingual speakers who spoke both Spanish and Catalan were excluded from the study to maintain the homogeneity of the language background. All of participants responded to a flyer advertisement, and their L2 profiles were carefully scrutinized via a telephone interview by means of background questionnaire. The study focused on experienced, active users, specifically individuals who regularly use English in their daily lives (Flege, Munro, & MacKay, 1995), and late L2 learners. These participants arrived in the UK after the age of 17 and had been residing there for at least a year, demonstrating stable L2 proficiency as outlined by Derwing, Munro, & Thomson (2008). The primary objective was to investigate the relative contributions of language experience, auditory acuity, and two memory systems to enhanced L2 proficiency.

In accordance with the screening criteria, we recruited participants with a range of chronological ages at the time of the study ( $M = 30$  years,  $Range = 19-47$ ), arrived in the UK after the age of 17 ( $M_{age} = 22$ ,  $Range = 17-33$ ) and had stayed for more than one year (LOR:  $M_{years} = 7.7$ ;  $Range = 1-20$ ). Recognizing that both the quantity and quality of experience matter in L2 learning (Flege, 2016), participants were carefully selected based on their average daily use of English in three distinct settings: school/work, home, and social life ( $M_{daily\ usage} = 70\%$ ,  $Range = 50-90\%$ ) (e.g., Flege & Liu, 2001; Flege et al., 1995). All participants reported significant experience in English-as-a-Foreign-Language education prior to their arrival in the UK ( $M_{years} = 7.53$  years;  $Range = 4-14$  years). Furthermore, none of the participants reported any prior hearing impairments that could potentially impact their perception and production of L2 speech.

To further examine the effects of experience, auditory processing, and memory in different phases of L2 speech learning, the 47 participants were further categorized into two subgroups as per their LOR profiles: (a) short-LOR ( $n = 20$ , LOR = 1-5 years) and (b) long-LOR ( $n = 27$ , LOR = 6-20 years). The distinction here corresponds to the way scholars have typically operationalized the two different phenomena in L2 learning, i.e., rate of learning vs ultimate attainment (see DeKesyer, 2013). In the rate of learning stage (LOR = 1-5 years), L2 learners are likely to demonstrate a dramatic improvement in their linguistic abilities. After years of immersion (LOR > 6 years), however, their L2 proficiency tends to become relatively stable and unchanged regardless of additional input and output opportunities. In the current study, the short-LOR group was assumed to represent the initial-to-mid state of L2 speech learning; and the long-LOR group was assumed to index their L2 attainment at the time of study.

The two-day experiment included auditory and cognitive abilities tests, administered via Zoom on the first day, and an IELTS speaking task on the second day. Throughout the entire duration of the test, the experimenter supervised the participants, providing assistance and addressing any questions they had. The tests focused on assessing auditory acuity, declarative memory, and procedural memory, with a 15-minute break provided between each test to ensure participants remained refreshed and attentive. Notably, no data was excluded from the study due to any interference or internet connection issues that may have arisen.

For the questionnaire materials, see **Supporting Information S1**.

## Auditory Processing Measures

In this current investigation, auditory processing was measured via a set of discrimination tasks developed in Kachlicka et al. (2019). The tasks were designed to assess participants' abilities to encode acoustic details of sounds. In these tasks, participants were required to listen to and differentiate between sounds using an AXB discrimination format.

Due to the COVID global pandemic, the task was mullet up and implemented on the Gorilla platform (Anwyl-Irvine et al., 2020), allowing remote access for participants. This approach was different from Kachlicka et al. (2019) wherein similar data was collected under face-to-face conditions.<sup>1</sup>

**Stimuli.** Two aspects of participants' auditory acuity were measured with three psychoacoustic tests, i.e., duration, formant, and pitch discrimination. Duration test scores were used for temporal acuity and formant and pitch test scores were used for spectral acuity. These stimuli were designed to not be perceived as speech by listeners. Each of the three subtests (formants, pitch, and duration) used the same set of stimuli, with the only difference being the target acoustic dimension.

The formant subtest consisted of 101 complex tones, with one standard stimulus and 100 comparison stimuli. Each stimulus had a duration of 500 ms and included 5-ms amplitude ramps at the beginning and end. The fundamental frequency was set at 100 Hz, with harmonics up to 3,000 Hz. Three formants were inserted at 500, 1,500, and 2,500 Hz using a parallel formant filter bank. The second formant of the standard stimulus was set at 1,500 Hz, while the comparison stimuli ranged from 1,502 to 1,700 Hz in 100 steps.

For the pitch and duration subtests, 101 four-harmonic complex tones were prepared. A 5-ms linear ramp was inserted at the beginning and end of each stimulus. In the pitch subtest, the fundamental frequency of the standard stimulus was set at 330 Hz, while the comparison stimuli ranged from 330.3 to 360 Hz in increments of 0.3 Hz. In the duration subtest, the length of the first amplitude ramp varied from 10 to 300 ms, with the standard stimulus set at 15 ms.

**Procedure.** During each trial, participants listened to three synthesized stimuli, with the second stimulus remaining constant. The first or last stimulus could vary, and participants were required to identify the sound that was different from the other two by pressing either the number "1" or "3" on a computer screen. Using Levitt's adaptive threshold procedure, the size of the difference between stimuli varied from trial to trial based on participants' performance. The tests started at the midpoint of the comparison stimuli and changed in increments of 10. An incorrect response decreased the difficulty by increasing the difference between stimuli by 10 steps, while three consecutive correct responses increased the difficulty by decreasing the difference between stimuli by 10 steps. After the first reversal, the step size changed from 10 to 5, and then from 5 to 1. The tests ended after 70 trials or eight reversals. Participants' auditory processing score was determined by the location of the final reversal, indicating how small a difference they could perceive between the standard and comparison stimuli. Lower scores indicated higher precision in perceptual acuity among participants. Although we did not specifically assess test-retest reliability in our study, Saito and Tierney (2022) have reported that the reliability of this test can be considered good.

For the audio materials used in the formant, pitch, and duration discrimination tests, see **Supporting Information S2**.

### **Declarative Memory**

LLAMA\_B is a cognitive task which has been commonly used in L2 aptitude research as a measure of explicit aptitude and declarative memory learning (Hamrick, 2015). LLAMA\_B is considered to be language-neutral, and it can be also referred to as rote learning or an associative memory (Li & Zhao, 2021). Drawing on the previous research (Hamrick, 2015) that looked at the role of individual differences in declarative and procedural memory learning, a similar 3.0 version of LLAMA\_B was used to measure participants'

---

<sup>1</sup> Adult participants likely perform the discrimination tasks similarly whether they engage in the tasks online or face-to-face modes (Saito & Tierney, 2024)

declarative memory learning ability (Meara, 2005). The test took approximately 10-15 minutes to complete. First, the participants were asked to learn the names of twenty unfamiliar objects within two minutes. Then they had to match the same but differently arranged objects with their names by clicking on them. This part of the test was not timed. After completing the test, the scores were calculated automatically by the software and ranged from 0-20. According to Cronbach's alpha analyses, the internal consistency of the declarative memory test ( $k = 20$ ) for 47 participants ( $n = 47$ ) yielded .953, which is far above the required benchmark ( $> .70$ ) in the field of applied linguistics (Larson-Hall, 2013).

For the materials used in LLAMA\_B, see

[https://www.lognostics.co.uk/tools/LLAMA\\_B/LLAMA\\_B\\_test.cgi](https://www.lognostics.co.uk/tools/LLAMA_B/LLAMA_B_test.cgi)

### **Procedural Memory**

The serial reaction time (SRT) task is widely utilized in L2 research to investigate implicit aptitude and procedural memory learning (Granena, 2013, 2019; Hamrick, 2015; Quam et al., 2018). This task is considered a reliable measure of implicit/procedural memory learning ability (Kaufman et al., 2010; Shanks, 2005). In the present study, we operationalized the deterministic version of the SRT task (Willingham, Nissen, & Bullemer, 1989), which has demonstrated high internal reliability across previous studies (Granena, 2019; Granena & Yilmaz, 2019) with a split-halves reliability of 0.79. In contrast, the probabilistic version of the SRT task (Kaufmann et al., 2010; Derwing & Munro, 2013; Granena, 2013; Suzuki & DeKeyser, 2015) exhibited lower reliabilities, ranging from 0.33 to 0.44. Specifically, in the deterministic SRT task, the experiment involves alternating blocks of patterned and random trials, with a greater number of patterned blocks compared to random blocks. In contrast, in probabilistic sequences, patterned trials are interspersed with random trials, and the signal-to-noise ratio varies across tasks (Oliveria et al., 2023).

Participants were required to download a free of charge Inquisit6 Lab software and STR task and were asked to complete the test on their computers. In line with previous research (e.g., Hamrick, 2015), the task included 6 blocks of 96 trials consisting of the first block, which presented a pseudorandom sequence (e.g.; V,B,N,M,V,N,B,V,M,B,M,N with the letters corresponding to the position of the red light appearing in one of four squares presented on the screen) and which was followed by four blocks of a patterned sequence (e.g. V,B,V,N,M,B,N,V,M,N,N,M), and the final sixth block of the repeated pseudorandom sequence. Participants were presented with 4 grey boxes in 4 possible screen positions and had to press a spatially corresponding response button on the keyboard as fast as possible once one of the boxes turned red. Following the procedure applied by Hamrick (2015), participants were asked to sit approximately 50 cm away from the computer screen to keep the stimulus angle set at the same threshold, which is approximately  $2.3^\circ \times 2.0^\circ$ . According to Destrebecqz and Cleeremans (2001), if intervals between stimuli (interval between response and presentation of next stimulus) are kept too long (e.g., 250-500 ms), explicit learning may take place. Thus, to ensure that participants were learning only implicitly, a response stimulus interval was set at 50 ms. Participants were not informed that they would encounter random and patterned sequences and/or that they were intended to learn a pattern.

Following the procedure operationalized in previous language aptitude studies (e.g., Hamrick, 2015; Linck et al., 2013), two scores were calculated based on the participants performance. First, the reaction times (RTs) were used as the indicator of a general processing speed. The lower RTs indicated a faster processing speed. However, lower RTs scores may be the result of the practice rather than of procedural memory learning. Hence, to factor out the effect of practice, the final pseudorandom block served as a control, and the RTs in this block are usually slower than on the final patterned block in healthy adults. Thus, the difference between RTs in pseudorandom and patterned block constituted a reactive facilitation score that indicated procedural memory learning. The higher the score, the better

the sequential learning. Finally, following Hamrick (2015), the reactive facilitation scores were submitted for further analyses. If the participant's average response accuracy in the fifth and sixth block was less than 70%, the participants' data had to be excluded (Linck et al., 2013). No data were excluded in the present study.

In line with previous research (Granena, 2013; Hamrick, 2015), the participants had to rate their familiarity with the sequence on a 6-point scale (1 = *I'm sure that this sequence was part of the test*; 6 = *I'm sure that this sequence was not part of the test*) to eliminate the possibility that explicit learning took place. None of the participants were able to remember the presented sequences. According to split-halves Spearman-Brown, the internal reliability of the SRT task yielded .952.

For the materials used in the SRT task, see <https://www.millisecond.com/products/inquisit6/weboverview.aspx>.

### Speaking Task

We adopted a procedure similar to the IELTS interview task in order to elicit more spontaneous and less controlled speech. Using the multifaceted framework of L2 oral ability (Crowther et al., 2015), we aimed to use seven different L2 oral ability measures so that we could tap into the four global aspects of L2 speech: (a) segmentals, (b) prosody (words stress and intonation), (c) fluency (rhythm and speech rate), and (d) lexicogrammar (vocabulary and grammar).

Although questionnaires, such as the LEAP-Q, are valuable tools for assessing language proficiency, they primarily focus on self-reported measures and may not capture the nuances of spontaneous spoken communication. By using the IELTS speaking task, we aimed to provide a more holistic evaluation of participants' language abilities, considering their proficiency in real-life communication scenarios. It allowed us for a more comprehensive evaluation of participants' oral communication abilities, including their ability to maintain a natural pace, use appropriate intonation, and pronounce words accurately being an important aspects of language fluency.

Participants joined scheduled Zoom meetings where the researchers shared the speaking task- IELTS interview topic, such as "What was the hardest and toughest challenge in your life?" and a set of follow-up questions, for example, "Why did you encounter this challenge?" The procedure followed by the researchers was based on the guidelines outlined in Saito, Trofimovich, and Isaacs (2017), which allowed participants one minute to prepare their response and 1-2 minutes to speak on the given topic. During the interview, participants' voice samples were recorded.

For speech materials, see **Supporting Information S3**.

### Raters

Previous research posited that multilingual raters tend to be more lenient than monolingual judges while conducting a speech assessment (Isaacs & Thomson, 2013). Therefore, we recruited three experienced raters who were native speakers of English ( $M_{age} = 34$  years; 3 females). All the participants raters had an ESL teaching experience which ranged from 1 to 10 years ( $M_{years\ of\ teaching} = 5.5$ ) and were trained for conducting a speaking assessment and teaching English phonology at a primary school. The NS raters were carefully selected and screened for their familiarity with Spanish-accented speech to allow for a more objective speech samples assessment. Consequently, the raters were asked to complete a questionnaire to report their familiarity with Spanish-accented English, which was  $M = 1,7$  on a 6-point scale (1 = *not at all*, 6 = *very familiar*). Finally, raters conducted an assessment of 47 speech samples by rating different dimensions of participants' L2 speech as presented in Table 1 included in Supporting Information.



For the raters' background questionnaire and materials, see **Supporting Information S4**.

### **Rating Procedure**

The judgements of 47 speech samples were conducted individually by native speaker raters in a quiet room on primary school premises. To ensure consistency, the raters were provided with training on each of the seven speech measures. Following the procedure outlined in Saito et al. (2017), first the researcher provided each rater with a brief explanation of the speech features to be assessed, specifically segmental accuracy. Next, the raters familiarized themselves with the IELTS task and the follow-up questions, and practiced the rating procedure by evaluating three speech samples that were not included in the main dataset (Saito, Trofimovich, & Isaacs, 2017).

For training scripts and descriptors, see **Supporting Information S5**

As for segmentals, word stress, intonation, rhythm, and speech rate, the assessment was audio-based which refers to the evaluation of these speech features solely based on the auditory information present in the speech samples. In other words, the raters relied on listening to the audio recordings of the speech samples rather than considering any visual or written cues. Only the first 30s of speech samples were used. Each speech sample was played once in the researcher's personal laptop. The raters were asked to assess each category on the 9-point Likert scale (1 = *not targetlike*, 9 = *targetlike*). Given that detailed lexical analysis requires longer samples (e.g., 3-4 minutes, see Saito et al. 2016), full length ( $M\ length = 3,3$ ) of participants' voice samples were transcribed orographically according to the procedure used by Crossley et al. (2014). Following this procedure, all audio transcripts were cleaned by removing mispronunciation errors (e.g., I went on the *sheep* cruise was transcribed as "I went on the cruise *ship*) and filled pauses (e.g., <sup>1</sup>uh, oh). The raters used a 9-point Likert scale to assess the lexical accuracy (1 = *many inappropriate words used*; 9 = *consistently appropriate vocabulary*) and the morphosyntactic accuracy (1 = *not targetlike*, 9 = *targetlike*). The inter-rater agreement for all linguistic concepts in the present study was high and reached for segmental errors: .972, word stress: .945, intonation: .935, rhythm: .933, speech rate: .935, lexical accuracy: .958, morphosyntactic accuracy: .965.

### **Results**

As summarized in Table 3, participants' L2 oral abilities, as well as perceptual-cognitive abilities, varied widely (for reference see *Table 3*). Since participants' declarative memory scores were significantly deviant from a normal distribution ( $p < .05$ ), their scores were transformed via the log10 function. Similar to the prior research (Saito et al., 2017), Pearson correlation analyses found participants' L2 oral abilities to be significantly correlated with each other ( $r = .705$  to  $.962$ ,  $p < .001$ ). In terms of the relationship between participants' auditory and memory abilities, a set of Spearman correlation analyses were conducted on the raw scores (an alpha set to .008 after Bonferroni corrections). Spectral acuity was significantly correlated with temporal acuity ( $r = .543$ ,  $p = .001$ ), suggesting that both measures tap into one broad category of auditory processing. Although there was a significant correlation between spectral acuity and declarative memory ( $r = -.409$ ,  $p = .004$ ), none of the other contexts reached statistical significance ( $r = -.346$  to  $-.112$ ,  $p = .017$  to  $.455$ ). The results here indicated that auditory processing, declarative memory, and procedural memory might be somewhat overlapping but essentially independent constructs (at least within the current study).

**Table 3**

*Descriptive Statistics of Participants' L2 Oral and Perceptual-Cognitive Abilities*

Variable	<i>M</i>	<i>SD</i>	95% CI	Normality <sup>a</sup>
<hr/>				

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

			Lower	Upper	<i>D</i>	<i>p</i>
<u>A. L2 oral abilities</u>						
Segmentals (9 points)	6.518	1.658	6.031	7.004	.169	.122
Word stress (9 points)	7.021	1.410	6.607	7.435	.120	.467
Intonation (9 points)	7.333	1.278	6.958	7.708	.138	.301
Rhythm (9 points)	6.589	1.510	6.145	7.032	.121	.461
Speech rate (9 points)	6.539	1.561	6.080	6.997	.193	.052
Lexical accuracy (9 points)	7.362	1.382	6.956	7.767	.148	.232
Morphosyntactic accuracy (9 points)	6.457	1.590	5.990	6.925	.150	.215
<u>B. Perceptual-cognitive abilities</u>						
Spectral acuity	.000	.833	-.245	-.245	.111	.571
Temporal acuity	.000	1.000	-.294	.293	.077	.920
Declarative memory	2.620	1.074	2.300	2.930	.194	.051
Procedural memory	133.77	75.715	111.54	156.00	.091	.797

*Note.* <sup>a</sup> for Kolmogorov-Smirnov Test. For spectral and temporal acuity lower scores are an indicator of a higher auditory sensitivity. For declarative and procedural memory- the higher the score the better sequential learning.

In the statistical analysis, our first objective was to examine the relationship between experience and age-related variables with participants' auditory processing (temporal and spectral acuity) and long-term memory (declarative and procedural memory) in order to answer the first research question (RQ1). Specifically, we aimed to investigate how auditory processing, procedural and declarative memory, and factors related to the experience of L2 learners (such as Length of Residence [LOR], Age on Arrival [AOA], Current L2 Use, and Foreign Language Education) interact with each other. This analysis allowed us to explore the complex interplay between these variables and provide insights into their influence on L2 learning. To this end, a set of multiple regression analyses were performed with participants' auditory and memory scores as dependent variables relative to four predictor variables: (a) age of arrival (AOA) in the UK (17-33 years), (b) length of residence (short vs long-LOR group), (c) current L2 use (%), and (d) foreign language experience.

As summarized in Table 4, none of the experience variables were significantly associated with participants' temporal and spectral acuity and declarative memory ( $p > .05$ ). In contrast, procedural memory was strongly tied to the LOR category (short- vs long-LOR;  $p = .003$ ). To follow up, an independent t-test was performed, to isolate and confirm the direct group-level difference in procedural memory scores between the two LOR groups without accounting for other predictor variables. This analysis found that the attainer group (LOR > 6 years) demonstrated larger procedural memory than the short-LOR group (LOR = 1-5 years) with large effects,  $t(45) = 3.257$ ,  $p = .003$ ,  $d = 0.942$ . In addition, a Pearson correlation analysis was conducted to find the relatively strong linear/continuous relationship between participants' individual differences in procedural memory and immersion profiles,  $r = .714$ ,  $p < .001$  (see *Figure 1*).

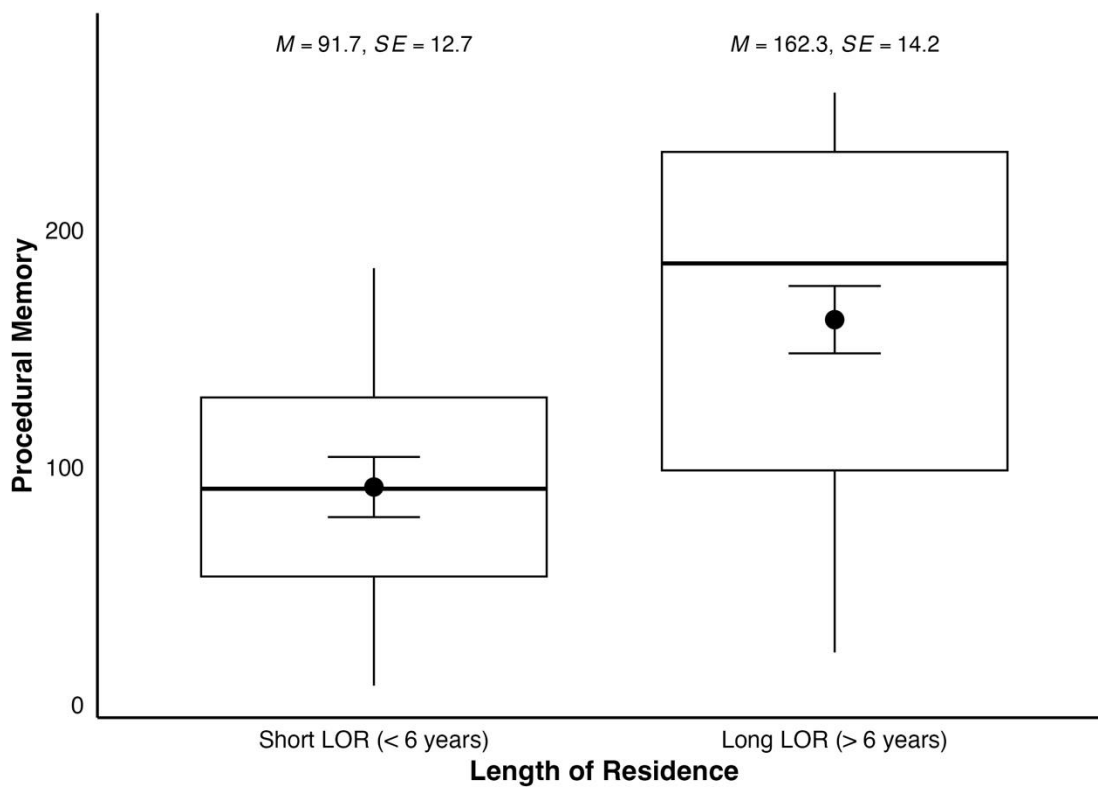
## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**Table 4**  
*Summary of Multiple Regression Analyses*

Predicted variables	Predictor variables	Standardized $\beta$	<i>SE</i>	<i>t</i>	<i>p</i>	<i>VIF</i> <sup>b</sup>	<i>R</i> <sup>2</sup>
Temporal acuity	AOA	-.131	.039	-.780	.440	1.197	.016
	LOR group <sup>a</sup>	.070	.322	.436	.665	1.106	
	Current L2 use	-.031	.011	-.169	.866	1.453	
	Foreign language education	.006	.074	.034	.973	1.324	
Spectral acuity	AOA	.074	.032	.444	.660	1.197	.018
	LOR group <sup>a</sup>	-.071	.268	-.439	.663	1.106	
	Current L2 use	-.046	.010	-.249	.804	1.453	
	Foreign language education	-.052	.062	-.294	.771	1.324	
Declarative memory	AOA	.052	.042	.308	.760	1.197	.046
	LOR group <sup>a</sup>	.088	.347	.547	.587	1.106	
	Current L2 use	.044	.012	.237	.814	1.453	
	Foreign language education	.013	.080	.076	.939	1.324	
Procedural memory	AOA	-.041	2.024	-.270	.789	1.197	.197
	LOR group <sup>a</sup>	.462	1.629	3.177	.003*	1.106	
	Current L2 use	.128	.588	.767	.448	1.453	
	Foreign language education	-.119	3.838	-.747	.459	1.324	

\**Note.* <sup>a</sup>for group categories (short- vs long-LOR); <sup>b</sup> for Variance Inflation Factor

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**FIGURE 1** Differences in Procedural Memory Between Short vs. Long Length-of-Residence Groups

\*Note. As for Procedural memory larger scores suggest higher ability.

### Perceptual, Cognitive, and Experiential Correlates of L2 Oral Ability

The final objective of the statistical analyses to address the RQ2 was to examine how both aptitude factors (temporal and spectral acuity, declarative and procedural memory) and bio factors (AOA, LOR, current L2 use, foreign language education) interacted to determine the outcomes of participants' L2 oral abilities. To this end, linear mixed-effects modelling analyses were conducted using seven different dimensions of participants' L2 speech proficiency scores i.e., vowel and consonant accuracy, words stress, intonation, rhythm, speech rate, and lexicogrammar (47 participants  $\times$  7 dimensions = 329 observations).

In Model 1, fixed effects included (a) temporal acuity, (b) spectral acuity, (c) declarative memory, (d) procedural memory, (e) age of acquisition, (f) length of residence, (g) current L2 use, and (h) foreign language education. Random effects included participants' ID (1-47) and composite scores of seven different dimensions of oral abilities (segmentals, word stress, intonation, rhythm, speech rate, lexical accuracy, and morphosyntactic accuracy) which were standardized and validated before entering the analysis. According to the Variance Inflation Factor, the eight fixed effects factors did not show clear evidence of multicollinearity (VIF = 1.065-1.511).

As summarized in Table 5, the model identified the significant main effects of spectral acuity ( $\beta = -.462$ ,  $t = -6.094$ ,  $p < .001$ ), procedural memory ( $\beta = .012$ ,  $t = 15.427$ ,  $p < .001$ ), and LOR ( $\beta = -.389$ ,  $t = -3.263$ ,  $p = .001$ ). The results indicated that those with more advanced L2 oral abilities were likely to have had not only a longer immersion experience (LOR > 6 years) but also higher levels of auditory precision and procedural memory. Effect size calculations indicated that spectral acuity demonstrated a medium-to-large effect ( $d = -.680$ ), procedural memory had a very large effect ( $d = 1.72$ ), and LOR exhibited a small-to-medium effect ( $d = -.360$ ).

Focusing on the three significant predictors (i.e., spectral acuity, procedural memory, and LOR), Model 2 was constructed to further examine the extent to which these factors interact to determine L2 oral abilities. Similar to Model 1, a factorial model yielded significant main effects of procedural memory ( $\beta = .012$ ,  $t = 13.962$ ,  $p < .001$ ,  $d = 1.56$ ) and LOR ( $\beta = -.662$ ,  $t = -2.986$ ,  $p = .003$ ,  $d = 0.330$ ). Whereas the main effects of spectral acuity became non-significant ( $\beta = -.289$ ,  $t = -1.362$ ,  $p = 0.174$ ), the interaction effects of auditory acuity and LOR reached statistical significance ( $\beta = -.597$ ,  $t = -2.339$ ,  $p = .020$ ,  $d = 0.260$ ).

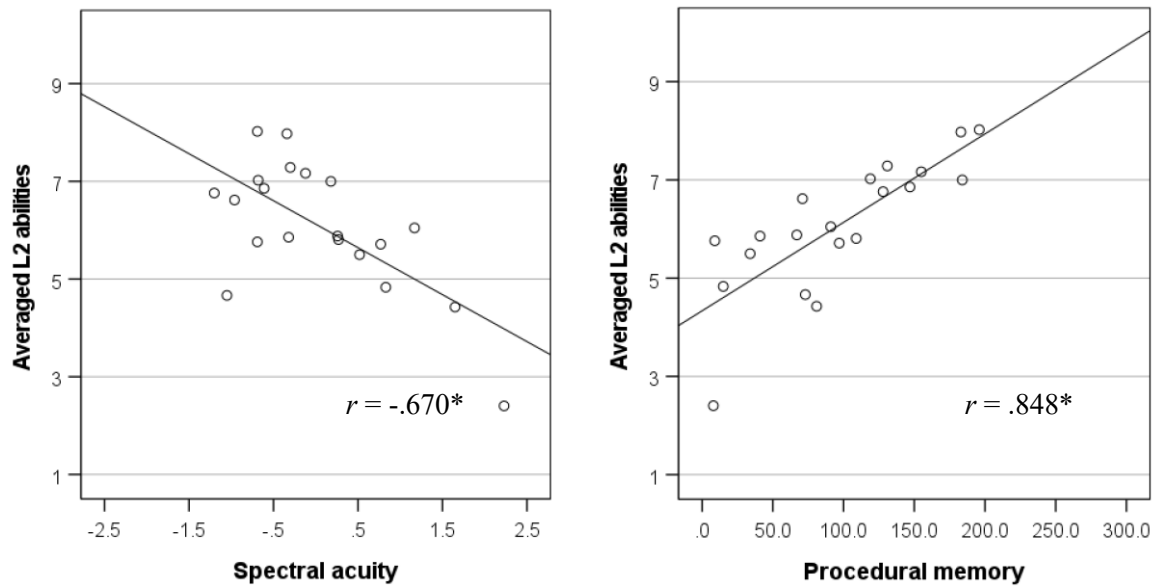
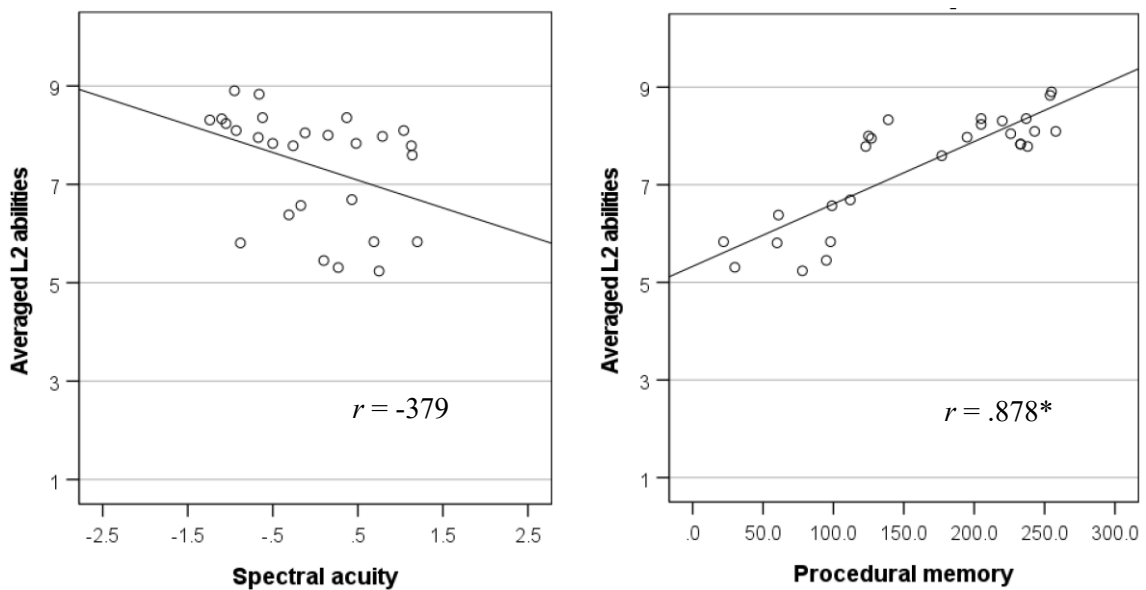
To examine the interaction between auditory processing and Length of Residence (LOR), we conducted two Pearson correlation analyses involving participants' spectral acuity and their averaged L2 oral abilities ( $n = 27$ ). This analysis allowed us to gain insight into the relationship between these variables and explore the potential impact of LOR on auditory processing in relation to L2 oral abilities. As visually summarized in Figure 2, the relationship between auditory processing and L2 oral abilities was strong among the Short-LOR group (LOR = 1-5 years;  $r = -.670$ ,  $p < .001$ ), but non-significant among the Long-LOR group (LOR > 6 years;  $r = -.379$ ,  $p = .051$ ). Interestingly, the correlation coefficients between procedural memory and averaged L2 oral abilities were comparable across the two groups ( $r = .848$  and  $.878$ ,  $p < .001$ ).

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**Table 5***Summary of Mixed-Effects Modelling Analyses of L2 Oral Abilities Relative to Eight Predictor Variables*

		Coefficient $\beta$	SE	t	p	95% CI	
						Lower	Upper
Model 1	Intercept	5.350	.4787	11.176	< .001*	4.408	6.292
	Temporal acuity	.077	.0592	1.301	.194	-.039	.193
	Spectral acuity	-.462	.0757	-6.094	< .001*	-.611	-.313
	Declarative memory	-.187	.1682	-1.111	.267	-.518	.144
	Procedural memory	.012	.0008	15.427	< .001*	.011	.014
	AOA	.010	.0130	.759	.449	-.016	.035
	LOR group						
	• Low (1-5 years)	-.389	.1192	-3.263	.001*	-.623	-.154
	• High (> 6 years)	.	.	.	.	.	.
	Current L2 use	-.001	.0038	-.371	.711	-.009	.006
Foreign language education	-.019	.0250	-.757	.450	-.068	.030	
Model 2	Intercept	5.408	.1577	34.295	< .000*	5.098	5.718
	Spectral acuity	-.289	.2124	-1.362	.174	-.707	.129
	Procedural memory	.012	.0009	13.962	< .001*	.011	.014
	LOR group						
	• Low (1-5 years)	-.662	.2215	-2.986	.003*	-1.097	-.226
	• High (> 6 years)1	.	.	.	.	.	.
	Spectral acuity $\times$ procedural memory	.001	.0012	.567	.571	-.002	.003
	Spectral acuity $\times$ LOR-Low	-.597	.2551	-2.339	.020*	-1.099	-.095
	Spectral acuity $\times$ LOR-High	.	.	.	.	.	.
	Procedural memory $\times$ LOR-Low	.003	.0017	1.524	.128	-.001	.006
Procedural memory $\times$ LOR-High	.	.	.	.	.	.	
Spectral acuity $\times$ Procedural memory $\times$ LOR-Low	.004	.0021	1.694	.091	-.001	.008	
Spectral acuity $\times$ Procedural memory $\times$ LOR-High	.	.	.	.	.	.	

## AUDITORY PROCESSING, MEMORY &amp; EXPERIENCE IN L2 SPEECH

**FIGURE 2** *Correlations Between Auditory Processing (Spectral Acuity), Procedural Memory, and L2 Oral Abilities as per Two Different Immersion Groups (Short vs Long)*<sup>3</sup>A. Short-LOR Group (n = 20 with 1-5 years of immersion experience)B. Long-LOR Group (n = 27 with 6-20 years of immersion experience)

*\*Note:* To calculate participants averaged L2 abilities (segmentals, word stress, intonation, rhythm, speech rate, lexical accuracy, and morphosyntactic accuracy) all scores were first standardized, summed up and divided by the number of abilities (in this case seven).

### Discussion

Focusing on a total of 47 Spanish speakers of English who were assumed to represent the initial-to-mid state of L2 speech learning ( $n = 20$  for 1-5 years of immersion) and the final state of L2 speech learning ( $n = 27$  for 6-20 years of immersion), the current study examined the complex relationship between auditory processing (temporal and spectral acuity), long-term memory (declarative and procedural), experience (LOR, AOA, current L2 use and foreign language education), and L2 oral abilities (segmentals, prosody, fluency, and lexicogrammar). With respect to the experience correlates of auditory processing and memory (RQ1), the results of multiple regression analyses showed that the model did not reveal any significant relationships between auditory processing or declarative memory and experience-related factors. In contrast, procedural memory yielded strong linear correlations with participants' profiles of the length of residence in the UK (1-20 years). The findings here suggest that both auditory processing and declarative memory can be considered as "aptitude", which is, by definition, a stable and unaffected by experience trait (Skehan, 2016).

Our argument is in line with the existing literature on the weak or nil relationship between auditory processing and L2 immersion experience (e.g., Saito et al., 2020a), although auditory processing can be different when analyses compare groups of individuals with substantially different experience profiles (e.g., see Bidelman, Gandour, & Krishnan, 2011, for tonal vs non-tonal language users; Krizman, Slater, Skoe, Marian, & Kraus, 2015, for simultaneous vs sequential bilinguals). Comparatively, very few studies have looked at the different types of memory profiles among L2 users with different experience and proficiency levels. There is some evidence showing that those with more L2 immersion experience tend to have a greater visuo-spatial and phonological short-term memory than monolinguals (e.g., Durand López, 2021; see also Anjomshoae, Nicoladis, & Gagnon, 2021) and that more advanced L2 users not only have ample immersion experience but also high-level cognitive capacities (e.g., Claussenius-Kalman, Hernandez, & Li, 2021; Linck et al., 2013). On a broad level, we echo Ullman's (2004, 2005) view that more advanced L2 learners may have a greater procedural memory in particular. At the initial stage of L2 learning, individuals may mainly adopt declarative memory to establish explicit linguistic knowledge. Throughout a long-term immersion experience, they gradually rely on procedural memory to access, reinforce, and automatize what they have declaratively learned. To our knowledge, however, little empirical work has been conducted on how the L2 immersion experience shapes procedural memory.

With respect to the relative weights of auditory processing, memory, experience, and L2 oral abilities, the results of mixed-effects modelling analyses showed that L2 outcomes can be determined by three different predictor variables, i.e., procedural memory, length of residence (short vs long), and auditory processing (spectral acuity). According to the follow-up analyses (see Model 2), it was shown that whereas procedural memory was equally predictive of L2 oral skills regardless of LOR profiles (short vs long), the link between auditory processing and proficiency was strong among those with a medium amount of immersion experience (LOR = 1-5 years) relative to those with an extensive amount of immersion experience (LOR = 6-20 years).

The findings presented here offer empirical backing for Skehan's (2016) aptitude-acquisition model, which emphasizes auditory processing as a significant influencer in the initial phases of L2 learning (Kachlicka et al., 2019) and underscores procedural memory as a crucial aptitude for achieving advanced levels of L2 proficiency later on (Linck et al., 2013). However, it is essential to approach these results with caution due to the limited sample size per LOR group. Future research directions could involve expanding the sample size to enhance the generalizability of the findings.



Interestingly, participants' declarative memory was not significantly correlated with L2 oral abilities (nor with any of the biographical backgrounds). This could be partially due to the fact that the current investigation highlighted that highly educated L2 learners with enough L2 linguistic knowledge (all participants had received years of foreign language education prior to immersion) may have already passed the very initial stage of L2 immersion (all participants had spent at least one year in the UK). In fact, previous research has shown that declarative memory plays a key role in the initial stage of language-focused, analytic L2 learning (Morgan-Short et al., 2014; Quam et al., 2018). The current study adds that L2 learners may not draw on declarative memory, especially when they are already immersed in L2 speaking environments (LOR > 1 year) where they regularly work on improving spontaneous oral L2 abilities with a view of achieving communication in various social settings (Granena, 2019).

Taken together, three tentative conclusions can be derived to disentangle the effects of auditory processing, memory, experience, and L2 speech learning. First, auditory processing and experience can make a separate contribution to L2 speech learning. Whilst all L2 learners can enhance L2 oral abilities as a function of increased experience, those with more precise spectral acuity may make the most of every input opportunity at the initial-to-mid stage of L2 speech learning (LOR 1-5 years), resulting in more advanced L2 oral proficiency in the long run. Second, the relationship between procedural memory, experience, and L2 oral abilities could be triangular in nature, the more immersion experience learners engage in, the more robust procedural memory they obtain and the more advanced, automatized, and nativelike L2 oral abilities they can use in the long run. Third, the acquisitional role of declarative memory remains open to further investigation as to the development of spontaneous L2 oral abilities in naturalistic settings.

### *References*

- Anjomshoae, F., Nicoladis, E., & Gagnon, R. (2021). Working memory capacity and structure in monolinguals and bilinguals. *Journal of Cognitive Psychology*, 33(4), 393-402. <https://doi.org/10.1080/20445911.2021.1923722>
- Anwyl-Irvine, A., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, 52(1), 388-407.
- Bidelman, G., Gandour, J., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience*, 23(2), 425-434. <http://dx.doi.org/10.1162/jocn.2009.21362>
- Brooks, P., Kwoka, N., & Kempe, V. (2017). Distributional effects and individual differences in L2 morphology learning. *Language Learning*, 67(1), 171-207. <http://dx.doi.org/10.1111/lang.12204>
- Chandrasekaran, B., Sampath, P. D., & Wong, P. C. M. (2010). Individual variability in cue-weighting and lexical tone learning. *The Journal of the Acoustical Society of America*, 128(1), 456-465. <https://doi.org/10.1121/1.3445785>
- Claussenius-Kalman, H., Hernandez, A.E., & Li, P. (2021). Expertise, ecosystem, and emergentism: Dynamic developmental bilingualism. *Brain and Language*, 222, Article 105013. <https://doi.org/10.1016/j.bandl.2021.105013>
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2014). Assessing lexical proficiency using analytic ratings: A case for collocation accuracy. *Applied Linguistics*, 36(5), 570-590. <http://dx.doi.org/10.1093/applin/amt056>

- Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2015). Does a speaking task affect second language comprehensibility? *The Modern Language Journal (Boulder, Colo.)*, 99(1), 80-95. <https://doi.org/10.1111/modl.12185>
- DeKeyser, R. (2013). Age effects in second language learning: Stepping stones toward better understanding. *Language Learning*, 63(1), 52-67. <https://doi.org/10.1111/j.1467-9922.2012.00737.x>
- Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, 63(2), 163-185. <https://doi.org/10.1111/lang.12000>
- Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29(3), 359-380. <https://doi.org/10.1093/applin/amm041>
- Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin and Review*, 8(2), 343-350. <https://doi.org/10.3758/bf03196171>
- Durand López, E. (2021). Morphological processing and individual frequency effects in L1 and L2 Spanish. *Lingua*, 257, Article 103093. <https://doi.org/10.1016/j.lingua.2021.103093>
- Eichenbaum, H. (2002). *The cognitive neuroscience of memory: An introduction*. New York: Oxford University Press.
- Flege, J. (2016). *The role of phonetic category formation in second language speech acquisition*. Plenary address delivered at New Sounds, Aarhus, Denmark.
- Flege, J., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23(4), 527-552.
- Flege, J., Munro, M., & MacKay, I. R. A. (1995). Factors affecting degree of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97(5), 3125-3134. <https://doi.org/10.1121/1.413041>
- Goswami, U., Fosker, T., Huss, M., Mead, N., & Szűcs, D. (2011). Rise time and formant transition duration in the discrimination of speech sounds: The Ba-Wa distinction in developmental dyslexia. *Developmental Science*, 14(1), 34-43. <https://doi.org/10.1111/j.1467-7687.2010.00955.x>
- Granena, G. (2013). Individual differences in sequence learning ability and SLA in early childhood and adulthood. *Language Learning*, 63(4), 665-703. <https://doi.org/10.1111/lang.12018>
- Granena, G. (2019). Cognitive aptitudes and L2 speaking proficiency. *Studies in Second Language Acquisition*, 41(2), 313-336. <https://doi.org/10.1017/S0272263118000256>
- Granena, G., & Long, M. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, 29(3), 311-343. <https://doi.org/10.1177%2F0267658312461497>
- Granena, G., & Yilmaz, Y. (2019). Corrective feedback and the role of implicit sequence-learning ability in L2 online performance. *Language Learning*, 69(1), 127-156. <https://doi.org/10.1111/lang.12319>
- Hamrick, P. (2015). Declarative and procedural memory abilities as individual differences in incidental language learning. *Learning and Individual Differences*, 44, 9-15. <https://doi.org/10.1016/j.lindif.2015.10.003>
- Hamrick, P., Lum, J. A. G., & Ullman, M. T. (2018). Child first language and adult second language are both tied to general-purpose learning systems. *Proceedings of the National Academy of Sciences*, 115(7), 1487-1492. <https://www.pnas.org/doi/full/10.1073/pnas.1713975115>

- Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, 10(2), 135-159. <https://doi.org/10.1080/15434303.2013.769545>
- Kachlicka, M., Saito, K., & Tierney, A. (2019). Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain and Language*, 192, 15-24. <https://doi.org/10.1016/j.bandl.2019.02.004>
- Kaufman, S., DeYoung, C., Gray, J., Jimenez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, 116(3), 321-340.
- Kempe, V., Thoresen, J., Kirk, N., Schaeffler, F., & Brooks, P. (2012). Individual differences in the discrimination of novel speech sounds: Effects of sex, temporal processing, musical and cognitive abilities. *PLOS ONE*, 7(11), Article e48623. <https://doi.org/10.1371/journal.pone.0048623>
- Kraus, N., & Banai, K. (2007). Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science*, 16(2), 105-110. <https://doi.org/10.1111/j.1467-8721.2007.00485>.
- Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience Letters*, 585, 48-53.
- Larson-Hall, J. (2016). *A guide to doing statistics in second language research using SPSS and R / Jenifer Larson-Hall* (2<sup>nd</sup> ed.). London: Routledge.
- Lengeris, A., & Hazan, V. (2010). The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native speakers of Greek. *The Journal of the Acoustical Society of America*, 128(6), 3757-3768. <https://pubs.aip.org/asa/jasa/article-abstract/128/6/3757/904260/The-effect-of-native-vowel-processing-ability-and?redirectedFrom=fulltext>
- Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467-477. <https://doi.org/10.1121/1.1912375>
- Li, S., & Zhao, H. (2021). The methodology of the research on language aptitude: A systematic review. *Annual Review of Applied Linguistics*, 41, 25-54. <https://doi.org/10.1017/S0267190520000136>
- Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., & Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, 63(3), 530-566. <https://doi.org/10.1111/lang.12011>
- Meara, P. (2005). *LLAMA language aptitude tests*. Swansea, UK: Lognostics.
- Morgan-Short, K., Faretta-Stutenberg, M., Brill-Schuetz, K., Carpenter, H., & Wong, P. (2014). Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism (Cambridge, England)*, 17(1), 56-72. <http://dx.doi.org/10.1017/S1366728912000715>
- Mora-Plaza, I., Saito, K., Suzukida, Y., Dewaele, J-M., & Tierney, A. (2022). *Tools for second language speech research and teaching*. <http://sla-speech-tools.com>. <https://doi.org/10.17616/R31NJNAX>
- Munro, M. J., Derwing, T. M., & Saito, K. (2024). Aspiring to aspirate: L2 acquisition of English word-initial /p/ over 10 years. *Language and Speech*. <https://doi.org/10.1177/00238309241264296>
- Muñoz, C. (2014). Contrasting effects of starting age and input on the oral performance of foreign language learners. *Applied Linguistics*, 35(4), 463-482. <https://doi.org/10.1093/applin/amu024>

- Nemeth, D., Janacsek, K., & Fiser, J. (2013). Age-dependent and coordinated shift in performance between implicit and explicit skill learning. *Frontiers in Computational Neuroscience*, 7, 147. <https://dx.doi.org/10.3389%2Ffncom.2013.00147>
- Poldrack, R.A. & Packard, M.G. (2003). Competition among multiple memory systems: Converging evidence from animal and human brain studies. *Neuropsychologia* 41(3), 245–51. [http://dx.doi.org/10.1016/S0028-3932\(02\)00157-4](http://dx.doi.org/10.1016/S0028-3932(02)00157-4)
- Quam, C., Wang, A., Maddox, W., Golisch, K., & Lotto, A. (2018). Procedural-memory, working-memory, and declarative-memory skills are each associated with dimensional integration in sound-category learning. *Frontiers in Psychology*, 9, 1828. <https://doi.org/10.3389/fpsyg.2018.01828>
- Rosen, S. & Manganari, E. (2001) Is there a relationship between speech and nonspeech auditory processing in children with dyslexia? *J of Speech, Language & Hearing Research* 44, 720-736.
- Saito, K. (2015). Experience effects on the development of late second language learners' oral proficiency. *Language Learning*, 65(3), 563–595. <https://doi.org/10.1111/lang.12120>
- Saito, K., Cui, H., Suzukida, Y., Dardon, D. E., Suzuki, Y., Jeong, H., & Tierney, A. (2022a). Does domain-general auditory processing uniquely explain the outcomes of second language speech acquisition, even once cognitive and demographic variables are accounted for? *Bilingualism: Language and Cognition*, 25(5), 856-868.
- Saito, K., Kachlicka, M., Suzukida, Y., Petrova, K., Lee, B. J., & Tierney, A. (2022b). Auditory precision hypothesis-L2: Dimension-specific relationships between auditory processing and second language segmental learning. *Cognition*, 229, 105236.
- Saito, K., Kachlicka, M., Suzukida, Y., Mora-Plaza, I., Ruan, Y., & Tierney, A. (2024). Auditory processing as perceptual, cognitive, and motoric abilities underlying successful second language acquisition: Interaction model. *Journal of Experimental Psychology: Human Perception and Performance*, 50(1), 119.
- Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, 69(3), 652-708. <https://doi.org/10.1111/lang.12345>
- Saito, Sun, H., Kachlicka, M., Robert, J., Nakata, T., & Tierney, A. (2020a). Domain-General Auditory Processing Explains Multiple Dimensions of Second Language Acquisition in Adulthood. *Studies in Second Language Acquisition*. <https://discovery.ucl.ac.uk/id/eprint/10106601/1/SSLA2020.pdf>
- Saito, K., Sun, H., & Tierney, A. (2020b). Domain-general auditory processing determines success in second language pronunciation learning in adulthood: A longitudinal study. *Applied Psycholinguistics*, 41(5), 1083 -1123. <https://doi.org/10.1017/S0142716420000491>
- Saito, K., Sun, H., Kachlicka, M., Alayo, J. R. C., Nakata, T., & Tierney, A. (2022c). Domain-general auditory processing explains multiple dimensions of L2 acquisition in adulthood. *Studies in Second Language Acquisition*, 44(1), 57-86.
- Saito, K., Suzukida, Y., Tran, M., & Tierney, A. (2021). Domain-general auditory processing partially explains second language speech learning in classroom settings: A review and generalization study. *Language Learning*, 71(3), 669-715. <https://doi.org/10.1017/S0142716420000491>
- Saito, K., & Tierney, A. (2022). Domain-general auditory processing as a conceptual and measurement framework for second language speech learning aptitude: A test-retest reliability study. *Studies in Second Language Acquisition*, 1–25. <https://doi.org/10.1017/S027226312200047X>

- Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*, 38(4), 439-462.  
<http://dx.doi.org/10.1093/applin/amv047>
- Segalowitz, N., & Freed, B. (2004). Context, contact and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition*, 26(2), 173-199.
- Shanks, D. R. (2005). Implicit learning. In K. Lamberts & R. Goldstone (Eds.), *Handbook of cognition* (pp. 202–220). Sage
- Skehan, P. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. In G. Granena, D. Jackson, & Y. Yilmaz (Eds.), *Cognitive individual differences in L2 processing and acquisition* (pp. 15-38). John Benjamins.
- Skoe, E., Krizman, J., Anderson, S., & Kraus, N. (2015). Stability and plasticity of auditory brainstem function across the lifespan. *Cerebral Cortex*, 25(6), 1415-1426.  
<https://doi.org/10.1093/cercor/bht311>
- Snowling, M. J., Gooch, D., McArthur, G., & Hulme, C. (2018). Language skills, but not frequency discrimination, predict reading skills in children at risk of dyslexia. *Psychological Science*, 29(8), 1270–1282.  
<https://doi.org/10.1177/0956797618763090>
- Sun, H., Saito, K., & Tierney, A. (2021). A longitudinal investigation of explicit and implicit auditory processing in L2 segmental and suprasegmental acquisition. *Studies in Second Language Acquisition*, 43(3), 551-573.
- Surprenant, A. M., & Watson, C. S. (2001). Individual differences in the processing of speech and nonspeech sounds by normal-hearing listeners. *The Journal of the Acoustical Society of America*, 110(4), 2085–2095.  
<https://doi.org/10.1121/1.1404973>
- Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, 65(4), 860-895.  
<http://dx.doi.org/10.1111/lang.12138>
- Tallal, P., Miller, S., & Fitch, R. H. (1993). Neurobiological basis of speech: A case for the preeminence of temporal processing. *Annals of the New York Academy of Sciences*, 682, 27-47. <https://nyaspubs.onlinelibrary.wiley.com/doi/10.1111/j.1749-6632.1993.tb22957.x>
- Tulving, E. (1993). What is episodic memory? *Current Directions in Psychological Science*, 2(3), 67–70.  
<https://journals.sagepub.com/doi/10.1111/1467-8721.ep10770899>
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1), 231–270.
- Ullman, M.T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory and practice* (pp. 141–78). Georgetown University Press.
- Ullman, M.T. (2020). The declarative/procedural model. A neurobiologically motivated theory of first and second language. In B. VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (3rd ed., pp. 128-161). Routledge.
- Ullman, M. T., & Lovelett, J. (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, 34(1), 39-65.  
<https://doi.org/10.1177%2F0267658316675195>

- Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4), 565-585.  
<https://doi.org/10.1017/S0142716407070312>
- Willingham, D. B., Nissen, M. J., & Bullemer, P. (1989). On the development of procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 1047–1060.
- Zheng, C., Saito, K., & Tierney, A. (2020). Successful second language pronunciation learning is linked to domain-general auditory processing rather than music aptitude. *Second language research*, 2020, 1-21.

### **Competing Interests Declaration**

Competing interests: The authors declare none.