Bilingualism: Language and Cognition

cambridge.org/bil

Does domain-general auditory processing uniquely explain the outcomes of second language speech acquisition, even once cognitive and demographic variables are accounted for?

Kazuya Saito, Haining Cui, Yui Suzukida, Diego Elisandro Dardon, Yuichi Suzuki, Hyeonjeong Jeong, Andrea Révész, Motoaki Sugiura and Adam Tierney

University College London, London, UK k.saito@ucl.ac.uk; Tohoku University, Sendai, Japan cui.haining.p5@dc.tohoku.ac.jp; University College London, London, UK yui.suzukida.18@ucl.ac.uk; Tohoku University, Sendai, Japan d.dardon7@gmail.com; Kanagawa University, Yokohama, Japan yuicchi0819@gmail.com; Tohoku University, Sendai, Japan jeong@tohoku.ac.jp; University College London, London, UK a.revesz@ucl.ac.uk; Tohoku University, Sendai, Japan motoaki.sugiura.d6@tohoku.ac.jp and Birkbeck, University of London, London, UK a.tierney@bbk.ac.uk

Research Article

Cite this article: Saito K, Cui H, Suzukida Y, Dardon DE, Suzuki Y, Jeong H, Révész A, Sugiura M, Tierney A (2022). 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**, 856–868. https://doi.org/10.1017/S1366728922000153

Received: 23 August 2021
Revised: 21 February 2022
Accepted: 21 February 2022
First published online: 27 April 2022

Keywords:
auditory processing; second language acquisition; speech; memory; executive functions

Address for correspondence:
Kazuya Saito
University College London,
Institute of Education
20 Bedford Way, United Kingdom WC1H 0AL
Email: k.saito@ucl.ac.uk

Abstract

Extending the paradigm in L1 acquisition, scholars have begun to investigate whether participants' domain-general ability to represent, encode, and integrate spectral and temporal dimensions of sounds (i.e., auditory processing) could be a potential determinant of the outcomes of post-pubertal L2 speech learning. The current study set out to test the hypothesis that auditory processing makes a UNIQUE contribution to L2 speech acquisition, for 70 Japanese classroom learners of English with different levels of L2 proficiency when biographical backgrounds (length of instruction and immersion) AND memory abilities (working, declarative, and procedural memory) are controlled for. Auditory processing loaded onto modality-general capacities to represent and incorporate anchor stimuli (relative to target stimuli) into long-term memory in an implicit fashion, but dissociated from explicit abilities to remember, associate, and elaborate sensory information. Auditory processing explained a small-to-medium amount of variance in L2 speech learning, even after the other potentially confounding variables were statistically factored out.

Introduction

Over the past 50 years, scholars have examined the foundations of successful post-pubertal second language (L2) speech learning, focusing on a wide range of populations within diverse learning contexts (naturalistic immersion vs. classrooms), stages (rate of learning vs. ultimate attainment), and dimensions (segmentals vs. suprasegmentals; accuracy vs. fluency; perception vs. production). There is evidence that learning outcomes are associated with a range of biographical variables, such as the quantity and quality of L2 input (Derwing & Munro, 2013), types of interlocutors (Flege & Liu, 2001), starting age of immersion (Abrahamsson & Hyltenstam, 2009), starting age of foreign language education (Muñoz, 2014), and learning contexts (Mora & Valls-Ferrer, 2012; for a comprehensive overview, see Saito, 2018). However, there is growing evidence that even if two learners spend the same amount of time practicing a target language in the same way within the same environment, their final outcomes will most likely be different (Doughty, 2019).

In the current study, we build on an emerging paradigm that domain-general auditory processing, which has been proposed to be a predictor of first language (L1) acquisition, plays a critical role in explaining some variance in post-pubertal L2 speech acquisition (Mueller, Friederici & Männel, 2012). Auditory processing is defined as one's capacity to precisely represent, discriminate, and categorize acoustic information. While the relationship between auditory processing and L1 acquisition remains open to discussion, growing evidence suggests that the effects of auditory processing could be relatively strong in the context of L2 acquisition in adulthood, especially in immersion contexts (e.g., Kachlicka, Saito & Tierney, 2019; Saito & Tierney, forthcoming; Sun, Saito & Tierney, 2021).

We hypothesize that the link between individual differences in auditory processing and language learning is not solely driven by a shared reliance on modality-general cognitive skills or a confounding influence of language experience but is at least partly driven by modality-

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

specific factors. To test this hypothesis, we examined data collected from 70 Japanese learners of English with different levels of L2 experience and proficiency to test the prediction that performance on auditory discrimination tests may help explain individual differences in L2 speech learning, even once language experience and memory capacity were accounted for. According to the major paradigms in L2 speech acquisition (e.g., Flege & Bohn, 2021 for Speech Learning Model), the capacities to hear the acoustic properties of new sounds anchors the development of perception and production abilities alike, suggesting that speech categorises are perception-based. In the current investigation, therefore, L2 speech proficiency was operationalized via an identification task (perceptually identifying contrasting phonemes in a minimal pair context).

Background

Domain-general auditory processing

In the field of cognitive psychology, much scholarly attention has been given to the perceptual and cognitive abilities underlying first language acquisition. Although the extent to which such networks are specific to language learning has remained open to debate (Campbell & Tyler, 2018), an influential view states that language-related processing involves the same neural systems responsible for general-purpose learning (see Hamrick, Lum & Ullman, 2018 for an overview). One such set of domain-general processes is AUDITORY PROCESSING, which collectively refers to a set of abilities related to the use of acoustic information, such as encoding, remembering, and integrating time and frequency characteristics of sounds. These abilities have been proposed to be a cornerstone of various learning behaviours related to speech, language, music, and emotion (Kraus & Banai, 2007).

In the audition-based account of language learning, auditory processing ability serves as "the gateway to spoken language" (Mueller et al., 2012, p. 15953) such that the detection and interpretation of acoustic information underlies every stage of phonetic, phonological, lexical, and morphosyntactic learning. For instance, toddlers analyse incoming acoustic input to detect segmental and suprasegmental patterns in a target language (Kuhl, 2004). Successful detection of these patterns can help children recognize word and sentence boundaries (Cutler & Butterfield, 1992), access and select contextually appropriate target words (Norris & McQueen, 2008), track syntactic structures (Marslen-Wilson, Tyler, Warren, Grenier & Lee, 1992), and fill in morphological details (Joanisse & Seidenberg, 1998). While there are other sources of lexical and morphosyntactic learning (e.g., visual input for reading), it is possible that auditory processing serves as a potential factor driving multiple stages of language acquisition (Tierney & Kraus, 2014).

Prior evidence suggests that individual differences in auditory processing precision may influence the rate and ultimate attainment of first language (L1) acquisition – i.e., the auditory-processing-deficit hypothesis (Goswami, 2015; Tallal, 2004; Wright, B. Bowen, R. & Zecker, 2000). Auditory impairments may slow down the speed of phonological, lexical, and morphosyntactic learning, eventually leading to a range of language problems. Supporting this hypothesis, individual differences in auditory processing abilities have been found to correlate with a range of L1 skills in children and adults, such as segmental and suprasegmental speech perception (Boets, Wouters, van Wieringen, De Smedt & Ghesquière, 2008; Cumming, Wilson & Goswami, 2015; Won, Tremblay, Clinard, Wright, Sagi & Svirsky, 2016), reading ability (Boets, Vandermosten, Poelmans, Luts, Wouters & Ghesquière, 2011; Casini, Pech-Georgel & Ziegler, 2018; Gibson, Hogben & Fletcher, 2006; Goswami, Wang, Cruz, Fosker, Mead & Huss, 2010; White-Schwoch, Carr, Thompson, Anderson, Nicol, Bradlow, Zecker & Kraus, 2015), phonological awareness (Moritz, Yampolsky, Papadelis, Thomson & Wolf, 2013; Peynircioglu, Durgunoglu & Oney-Kusefoglu, 2002), phonological memory (Tierney, White-Schwoch, MacLean & Kraus, 2017), syntax processing (Gordon, Shivers, Wieland, Kotz, Yoder & McAuley, 2015) and the incidence of language impairment (Corriveau & Goswami, 2009; Haake, Kob, Willmes & Domahs, 2013; McArthur & Bishop, 2004, 2005). Auditory processing measures have been suggested to be a diagnostic tool for dyslexia (Hornickel & Kraus, 2013) and other language-related disorders (Russo, Skoe, Trommer, Nicol, Zecker, Bradlow & Kraus, 2008; but see Rosen, 2003 for further discussion on the CAUSALITY of auditory processing and language acquisition).

Roles of auditory perception in L2 speech acquisition

Extending this line of thought, scholars have argued that the acquisitional role of auditory processing might be germane to L2 learning as well; in fact, it might be even more integral than it is for L1 acquisition, precisely because of the QUANTITATIVE and QUALITATIVE differences between L1 and L2 learning experience. Although children with developmental language delays (such as dyslexia) are more likely to initially have auditory deficits, auditory-based difficulties with speech and language perception may be mitigated in the long run through the extensive practice and experience with spoken language available in the context of L1 acquisition. Comparatively, adult L2 learners' access to target language input is limited in quantity. Many learners start studying a target language at a relatively later age through foreign language education (several hours of form-focused instruction per week) (Muñoz, 2014), and have far fewer opportunities to interact with different types of native interlocutors in various social settings even under immersion conditions (Derwing & Munro, 2013). This lack of interactive, variable, and sufficient input opportunities jeopardizes the development of compensatory strategies, thereby making any auditory deficit potentially even more problematic over the course of L2 learning (Saito, Sun & Tierney, 2020b).

Moreover, one's degree of auditory processing precision may be a more consequential skill in L2 speech acquisition (compared to L1 speech acquisition), due to the relatively demanding nature of acoustic analysis, integration, and adaptation in this context. Unlike L1 acquisition, wherein learning takes place in an auditory system relatively unaffected by language experience, adult L2 learners are required to encode individual acoustic dimensions of new sounds through the lens of the already-established, automatized acoustic representations developed through exposure to their L1 (Flege & Bohn, 2021). From a theoretical standpoint, auditory analyses during L2 speech learning are subject to the complex interaction between new and previous language experience (e.g., McAllister, Flege, J. & Piske, 2002 for feature hypothesis). Not only do learners restructure their L1-specific weights of acoustic cues (e.g., Chinese speakers learning to attend to both pitch and duration for perceiving L2 English prosody; Jasmin, Sun & Tierney, 2021), but they also detect new spectro-temporal patterns that they would otherwise not use during L1 processing (e.g., Japanese speakers attending to variations in third formants for perceiving L2 English [r] and [l]; Iverson, Kuhl, Akahane-

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

Yamada, Diesch, Tohkura, Kettermann & Siebert, 2003). When L2 learners poorly perceive certain auditory dimensions, they may show difficulty in learning new sounds, as they continue to filter out L2 input based on their L1 acoustic representations (Perrachione, T. Lee, Ha, L. Y. & Wong, 2011).

Previous research has shown that adult L2 learners with greater auditory sensitivity can better perceive novel phonetic and phonological contrasts in a language that they have never heard (Jost, Eberhard-Moscicka, Pleisch, Heusser, Brandeis, Zevin & Maurer, 2015; Kempe, Bublitz & Brooks, 2015); and that their degree of auditory processing precision predicts the extent to which adult L2 learners can improve their perception of novel and foreign language sounds when they receive intensive training (Lengeris & Hazan, 2010; Wong & Perrachione, 2007). Building on these laboratory studies, Omote, Jasmin, and Tierney (2017) more recently examined the relationship between auditory processing and L2 English perception proficiency among 25 adult Japanese speakers with varying amounts of naturalistic IMMERSION experience in the UK. The results showed that auditory processing (i.e., neural encoding of sound, as assessed using the frequency-following-response) was strongly predictive of L2 perception attainment, even when all other biographical variables were statistically controlled, including length of immersion and age of acquisition.

The emerging findings in support of the relatively strong connection between auditory processing and post-pubertal L2 speech acquisition have been replicated with other groups of L2 learners with diverse L1 backgrounds (e.g., Chinese, Polish, Spanish; Saito, Sun, Kachlicka, Robert, Nakata & Tierney, 2021b) across different dimensions of L2 learning (e.g., Kachlicka et al., 2019 and Saito, Macmillan, Kroeger, Magne, Takizawa, Kachlicka & Tierney, 2022 for lexicogrammar learning; Saito, Sun & Tierney, 2019; Saito, Kachlicka, Sun & Tierney, 2020a; Saito, Suzukida, Tran & Tierney, 2021a for speech production) from both cross-sectional and longitudinal perspectives (Saito et al., 2020b; Sun et al., 2021). On the whole, it seems that the outcomes of L2 speech learning could be EQUALLY affected by factors related to auditory processing and biographical backgrounds. Whereas all L2 learners demonstrate significant improvement when they engage in naturalistic immersion, certain L2 learners with greater auditory sensitivity are likely to make more of every input opportunity, converting more input into intake, and leading to more advanced L2 speech proficiency in the long run (for a comprehensive review, Saito & Tierney, forthcoming).

Motivation for the current study

Though promising, these initial findings raise a number of issues that need to be further investigated in order to provide a more comprehensive picture of the mechanisms underlying the relationship between auditory processing and adult L2 speech learning. In particular, the current study is designed to explore the complex relationship between perceptual, cognitive, and biographical individual differences, and their composite impact on L2 speech perception and production behaviours.

Although it is clear that there is a link between precise auditory perception and successful L2 speech acquisition, it remains unclear whether this link is driven by a confounding relationship with other cognitive capacities. Specifically, certain scholars have proposed that the link between auditory processing and language learning reflects a shared reliance of both tasks on short- and long-term memory abilities (Ahissar, Lubin, Putter-Katz & Banai, 2006). According to this position, L1 learners with dyslexia are more likely to have auditory processing problems, because they have poorer implicit memory capacities as manifest in less prolonged neural adaptation (i.e., shorter windows of temporal integration), suggesting that memory problems may drive both the reading difficulties and the auditory processing difficulties (Jaffe-Dax, Frenkel & Ahissar, 2017).

Behavioural tasks for measuring auditory perception (e.g., formant, pitch, and duration discrimination; for details, see below) may tap into a set of modality-general executive skills as well. Sound discrimination tasks are highly repetitive and abstract, making it difficult for participants to maintain auditory information in short-term memory (Bidelman, Gandour & Krishnan, 2009). During multiple exposures to the same audio stimuli, those with greater memory capacities could maintain an accurate representation for a longer period of time, making more sensory information available for acoustic analyses (Zhang, Moore, Guiraud, Molloy, Yan & Amitay, 2016). In addition, sound discrimination tests may trigger implicit statistical learning of the distribution of stimuli across trials, such that the prior stimulus distribution is combined with the representation of each incoming stimulus (Raviv, Ahissar & Loewenstein, 2012). Thus, greater implicit and statistical learning ability may help learners to extract the prototypical percept from multiple exposures more easily and reliably, which may in turn help enhance accurate auditory perception, especially when the current sensory stimulus includes noise.

Therefore, it remains to be seen whether, to what degree, and how auditory processing is UNIQUELY linked to the degree of success in L2 speech learning, even once individual differences in explicit and implicit memory are accounted for. If auditory processing is a somewhat INDEPENDENT construct (with only partial overlap with other learner-external and -internal factors), the link between auditory perception and proficiency should remain significant, even when the effects of the executive function and biographical factors are controlled for. Addressing this question will shed light on whether the evidence for the link between auditory processing and L2 speech acquisition supports the theoretical claim that perceptual acuity is a bottleneck of language learning throughout one's lifespan (Cumming et al., 2015).

To date, many studies have demonstrated a significant role for phonological short-term memory in language learning, when learners are exposed to new sounds and words that they have never heard before (Baddeley, 1993; Gathercole, Frankish, Pickering & Peaker, 1999; Papagno, Valentine & Baddeley, 1991; Reiterer, Hu, Erb, Rota, Nardo, Grodd, Winkler & Ackermann, 2011); however, such short-term memory abilities may be irrelevant, as learners engage in more extensive and intensive practice in order to ACQUIRE the target language (Hu, Ackermann, Martin, Erb, Winkler & Reiterer, 2013). Though few, some scholars have begun to investigate the mediating roles of cognitive abilities (i.e., short- and long-term memory capacities) in L2 phonological ACQUISITION. For example, Darcy, Park and Yang's (2015) study explored the relationship between a total of 30 experienced and inexperienced Korean learners' L2 English vowel proficiency and working memory abilities (see also Silbert, Smith, Jackson, Campbell, Hughes & Tare, 2015 for declarative memory).

Our proposed study was the first attempt to test the hypothesis that perceptual factors predict independent variance in L2 speech acquisition, even once cognitive factors are accounted for. If we assume that auditory processing could be a somewhat

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

INDEPENDENT construct (with only partial overlap with other learner-external and -internal factors), the link between auditory perception and proficiency should remain significant, even when the effects of the executive and biographical factors are controlled for. The findings here were expected to shed light on whether the evidence for the link between auditory processing and L2 (and L1) speech acquisition supports the theoretical claim that perceptual acuity is a bottleneck of language learning throughout one's lifespan (Cumming et al., 2015).

Current study

In the current investigation, a total of 70 young adult Japanese learners of English with a wide range of L2 experience and proficiency levels were recruited at a university in Japan. As summarized in Table 1, participants completed a battery of tasks designed to tap into auditory processing (formant, pitch, and duration discrimination), short-term memory (phonological memory, working memory), and long-term memory (declarative and procedural memory). These memory abilities could also be categorized as explicit versus implicit: the three tasks (phonological short-term memory, working memory, declarative memory) inevitably require participants' awareness (Li, 2016), but the last task (procedural memory) is hypothesized to be free of such awareness (Linck, Hughes, Campbell, Silbert, Tare, Jackson, Smith, Bunting & Doughty, 2013).

Afterwards, the participants took part in an individual interview to report their current and past experience in L2 and music learning. Finally, their perceptual, cognitive, and biographical profiles were linked to their L2 English speech perception and production proficiency, elicited from forced-choice identification and spontaneous picture description tasks. The following research questions and predictions were formulated:

R1: To what degree are participants' individual differences in auditory perception and memory functions inter-related?

In accordance with the literature on L1 acquisition, some overlaps were thought to exist between individual differences in auditory processing and short- and long-term memory functions. This is arguably because auditory processing draws on or/and co-functions with short-term memory skills (Sharma, Purdy & Kelly, 2009) and/or because auditory processing tasks per se (i.e., sound discrimination) may require participants to access statistical learning of the association, distribution, and probability of the anchor vs. target stimuli (Raviv et al., 2012).

R2: Does auditory processing explain independent variance in L2 speech perception and production proficiency, even once memory capacities and language experience are accounted for?

It has been shown that those with more advanced L2 speech proficiency are likely to have not only more extensive and intensive L2 training experience (Muñoz, 2014), but also more precise auditory perception (Kachlicka et al., 2019). In light of Plonsky and Oswald's (2014) field-specific guidelines, the size of the audition and acquisition link has been generally medium-to-large (r = .4-.6). Whereas research on the role of working memory in L2 phonology is still limited (cf. Darcy et al., 2015), Linck, Osthus, Koeth and Bunting's (2014) meta-analysis showed that the relationship between working memory and other areas of L2 acquisition (e.g., grammar) appears to be relatively small (r = .25). Some overlap may exist between individual differences in auditory processing and short- and long-term memory capacity (Sharma et al., 2009). Nevertheless, auditory processing was hypothesized to primarily reflect the degree of internal noise in the encoding of acoustic information in the early auditory system. Our prediction was that the relationship between auditory processing and L2 speech perception would remain significant, even when memory span and language experience were accounted for.

Table 1. Summary of Auditory Processing and Cognitive Measures

Constructs	Subconstructs	Measures
Auditory processing	Perceptual acuity	Combination of formant, pitch, and duration discrimination
Short-term memory	Phonological short-term memory (explicit)	Non-word repetition
	Working memory (explicit)	Operation span
Long-term memory	Declarative memory (explicit)	LLAMA-B
	Procedural memory (implicit)	Serial reaction time

Method

Participants

To examine the role of auditory processing in various phases of L2 speech acquisition among post-pubertal L2 learners with widely different biographical backgrounds, a total of 70 college-level Japanese learners of English were recruited at a relatively large public university in Japan. Given that auditory processing is somewhat related to chronological age (Skoe, Krizman, Anderson & Kraus, 2015), we recruited a group of young adults within a restricted age range (18–26 years). Given that the presence/absence of naturalistic immersion experience plays a key role in determining the process and product of L2 speech acquisition (Derwing & Munro, 2013; Muñoz, 2014), our participants comprised (a) 51 inexperienced participants who had studied L2 English only through foreign language education without any experience abroad; and (b) 19 experienced participants who had study-abroad experience (i.e., more than one month of stay in English speaking countries). All of them were enrolled in undergraduate or graduate-level courses at the time of the project.

Design

The Japanese participants engaged in a series of tasks to investigate: their language experience and backgrounds (5 minutes), auditory processing abilities (10 minutes), working and long-term memory capacities (40 minutes), and L2 English speech proficiency (10 minutes). To call for participants, an electronic flyer was distributed across the university. Interested participants contacted an investigator and then joined an individual meeting. To avoid any unwanted confusion among the participants, all the tasks were delivered in Japanese.

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

Experience factors

Focusing on a similar L2 population (university-level English-as-a-Foreign-Language students), precursor research has examined what kinds of experience variables could be linked to the rate of L2 speech acquisition in classroom settings. The results showed that the outcomes of advanced L2 speech learning could be linked to not only the length of EFL instruction, but also to participants' unique L2 learning experience profiles, such as study abroad (e.g., Muñoz, 2014) and pronunciation training (e.g., Saito & Hanzawa, 2016). Thus, the three variables (length of instruction, study abroad, pronunciation instruction) were taken into consideration in the current study. By modifying the EFL experience questionnaire (Saito & Hanzawa, 2016), each participant was interviewed to collect and code the following information[1]:

Length of instruction

There is extensive prior evidence that adult L2 learners' speech proficiency is strongly associated with the length of foreign language education (Muñoz, 2014). The participating college students likely have different onsets of foreign language education (Grades 1-6), resulting in different amounts of instructed L2 speech learning experience. The length of foreign language education was coded as a continuous variable ($M = 11.1$ years; $SD = 3.48$; *Range* = 6–19).

Specific experience

To further delve into the quality of L2 learning experience, the two unique experience profiles of participants were coded in terms of the presence of English pronunciation training ($n = 36$ for yes, 34 for no) and study-abroad experience ($n = 19$ for yes, 51 for no; Muñoz, 2014). As in Saito and Hanzawa (2016), pronunciation training was defined as provision of explicit instruction on the perceptual and/or articulatory characteristics of L2 English phonemes which are known to be difficult for a particular group of L1 Japanese speakers (e.g., English [r] and [l]). As reported in many educational reports (e.g., Saito, 2014), few teachers receive training on pronunciation training, and pronunciation instruction is ignored in Japanese EFL classrooms. Japanese EFL syllabus typically highlight the rote memorization of vocabulary items and grammar rules followed by reading and writing exercise without much focus on listening and speaking (Nishino & Watanabe, 2008).

Auditory processing measures

For the purpose of comparison with the existing literature on L1 and L2 acquisition (e.g., Kachlicka et al., 2019), participants' acuity to three different dimensions of sounds (formant, pitch, and duration) was to be measured via three different auditory discrimination tasks. For each task, the participant listened to three sounds, two of which were identical, and the third of which was different; their task was to identify which sound was the different sound. Either the first or third sound was always different. For each test, a series of 100 synthesized stimuli were prepared (500 ms in length) along a target acoustic dimension via custom MATLAB scripts. In the precursor project, the test-retest reliability of the test format was examined and confirmed among a range of L1 and L2 speakers ($r = .701$, $p < .001$; Saito, Sun & Tierney, 2020c).

Stimulus

In the formant discrimination test, the fundamental frequency was set at 100 Hz, and 30 harmonics were created up to 3000 Hz. A parallel formant filter bank was used to impose three formants at 500 Hz, 1500 Hz, and 2500 Hz (Smith, 2007). For the comparison stimulus, the target dimension (i.e., F2) varied between 1502 Hz and 1700 Hz with a step of 2 Hz. In the pitch and duration discrimination tests, the stimuli were constructed by modifying a baseline four-harmonic complex tone with a fundamental frequency of 330 Hz, a duration of 500 ms, and a 15 ms linear amplitude ramp at the beginning and end. For the pitch discrimination test, the fundamental frequency of the baseline stimulus was set at 330 Hz, while that of the comparison stimulus ranged from 330.3 to 360 Hz with a step size of 0.3 Hz. For the duration discrimination test, the duration of the baseline stimulus was set at 250 ms, while that of the comparison stimulus ranged from 252.5-500 ms with a step size of 2.5 ms. All the audio files are currently available for reviewers to access for the purpose of thorough peer review (see Supporting Information A, Supplementary Materials).

Procedure

Three different tones were presented with an inter-stimulus interval of 0.5 s. Upon hearing them, the participants chose which of the three tones differed from the other two by pressing the number "1" or "3." Using Levitt's (1971) adaptive threshold procedure, the level of difficulty changed from trial to trial based on participants' performance. Each test started with the standard stimulus (Level 50). When three correct responses were made in a row, the difference became smaller by a degree of 10 steps (more difficult). When their response was incorrect, the difference became wider by a degree of 10 steps (easier).

The step size changed when the direction of difficulty between trials reversed – i.e., when an increase in acoustic difference (easier) was followed by a decrease (more difficult), or vice versa. After the first reversal, the step size decreased (more difficult) from 10 to 5, and then after the second reversal the step size decreased further from 5 to 1. The tests stopped either after 70 trials or eight reversals. To calculate participants' auditory processing scores, the stimulus levels of the reversals from the third reversal until the end of the block were averaged. Performance across all three tests (pitch, duration, and formant discrimination) was standardized and averaged (after transforming scores on each test to z-scores) to form a composite measure of auditory processing.

Cognitive ability measures

Cognitive abilities are defined as the capacity to transform, remember, use, and elaborate sensory information. In the current study, we assessed participants' cognitive abilities with a focus on short- and long-term memory under single and/or dual task

[1]Although the questionnaire elicited participants' starting age of EFL instruction, not surprisingly, the variables were not significantly related to their L2 outcomes ($p > .05$). To date, there was little empirical evidence in support of the role of age of learning in classroom L2 learning (see Muñoz, 2014 for a comprehensive overview on the length and timing of EFL instruction). Age of acquisition has been found as a strong predictor when research has focused on naturalistic settings wherein learners can receive ample high-quality input. In classroom contexts, however, input is limited to several hours of language-focused instruction. Additionally, age of learning was excluded because it was highly correlated with length of learning ($r = -.855$, $p < .001$). The negative correlation was not surprising as those with earlier age of learning tended to spend longer time in EFL education. Including both variables could have inflated the multicollinearity problems in the regression analyses.

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

conditions. We did so in accordance with two major theoretical accounts of cognitive individual differences in L2 acquisition – i.e., working memory (Linck et al., 2014) and declarative-procedural memory (Ullman & Lovelett, 2018). Working memory refers to a set of cognitive functions responsible for the control, regulation, and storage of sensory information (e.g., Conway, Jarrold, Kane, Miyake & Towse, 2007). In Baddeley's (2003) influential model, the construct of working memory is divided into two subsystems: (a) a storage-based system which maintains short-term memory; and (b) an executive control system which manages information between short- and long-term memory stores.

In terms of long-term memory, one influential paradigm in cognitive psychology is Ullman's declarative and procedural memory model (Ullman, 2004). Declarative memory is thought to be primarily responsible for the explicit learning of episodic and semantic knowledge and associations (Tulving, 1993), and is directly relevant to the early phase of lexical learning (memorizing the sound/meaning of words, irregular morphological forms, and idioms; Ullman & Lovelett, 2018). Procedural memory is a key component of implicit and statistical learning and the control of sensorimotor and cognitive skills, such as sequences, rules, and categorization (Knowlton, Mangels & Squire, 1996). In the context of L2 learning, it has been suggested that procedural memory is instrumental to the later stages of grammar language learning (acquiring rule-based morphology, syntax and phonology; Ullman & Lovelett, 2018).

As operationalized in Linck et al. (2013), four measures were used to tap into participants' cognitive abilities: non-word repetition to assess phonological short-term memory (e.g., Révész, 2012); operation span to test working memory (Turner & Engle, 1989); paired associates to measure declarative memory (Morgan-Short, Faretta-Stutenberg, Brill-Schuetz, Carpenter & Wong, 2014); and serial reaction time for procedural memory (Kaufman, DeYoung, Gray, Jiménez, Brown & Mackintosh, 2010).

Phonological short-term memory

Participants completed a Japanese version of the non-word repetition task (Yamaguchi & Shimizu, 2011), wherein they were asked to remember and repeat random sequences of non-words that follow Japanese phonotactic rules. The test consisted of 20 trials; there were five trials for each set consisting of one, two, three, or four nonwords. For instance, in a four nonword trial, participants heard and repeated four nonwords (e.g., *tesaya, nimika, kekayu, yuteka*). Credit was given for each correctly repeated word, and the total number of successful repetitions (with 50 as the maximum) was used as a PSTM score, indicating how well participants store phonological information and articulate the words. Cronbach alpha was .77.

Complex working memory

Participants completed an automated operation span test, wherein they were asked to remember a sequence of alphabetical letters while solving math problems under dual task conditions (Unsworth, Heitz, Schrock & Engle, 2005). For each trial, participants first answered whether the solution of an equation was correct (e.g., (1*2) + 1 = 3), and then remembered a letter of the alphabet. After one trial was completed, participants selected the letters in the order in which they were presented. There were 15 sets of letters ranging from 3 to 7, and the total number of letters is 75 (3, 3, 3, 4, 4, 4, 5, 5, 5, 6, 6, 6, 7, 7, 7). The possible maximum score is 75. The Cronbach alpha reliability of the test was .721.

Declarative memory

Participants completed the LLAMA-B test (Meara, 2005), which has been used as a behavioural measure of declarative memory (Morgan-Short et al., 2014). In this paired associates task, participants were asked to remember the combinations of names and shapes of 20 objects. The participants first learned as many words as possible within two minutes by clicking on images to display the names of various objects. In the following testing phase, they were asked to correctly link the names of randomly chosen objects with the correct picture. The possible maximum score was 20. Cronbach alpha was .75.

Procedural memory

Participants completed a statistical serial reaction time task, which has been used as a behavioral measure of procedural memory (Kaufman et al., 2010). In this task, participants saw a dot appearing at one of four locations on the computer screen and responded to it as quickly and accurately as possible by pressing the corresponding key. Unbeknownst to participants, the sequence of stimuli was generated by a probabilistic rule: 85% of the sequences followed the rule (probable sequence), whereas the other 15% of the sequences were generated by another rule (improbable sequence). There were eight blocks, and each block consisted of 120 trials, with 960 trials in total. The task was scored by subtracting the mean RTs to the probable sequences from those to the improbable sequences, which reflects the amount of learning. Split half reliability was .538.

L2 speech perception measures

To measure participants' L2 speech perception proficiency, they completed a two-alternative forced-choice identification test, developed by Slevc and Miyake (2006). The test format comprised a total of 51 minimal pairs (e.g., late vs. rate; filling vs. feeling), wherein participants were asked to discriminate a range of consonants (e.g., [r] vs. [l]) and vowels (e.g., [ɪ] vs. [i]) that are particularly difficult for Japanese learners of English.

There are sizeable phonological differences between Japanese and General American English vowels (English: 15; Japanese: 5) and consonants (English: 24; Japanese: 14). Thus, Japanese speakers have difficulty noticing, learning, and processing the English phonological contrasts absent in the L1 phonetic systems. Given these differences between Japanese and English phonology, Saito (2014) provided a comprehensive overview on why these English contrasts were problematic for Japanese listeners:

- **Vowels (Monophthongs):** Japanese listeners perceive low front vowel [æ], central mid vowel [ʌ], and low back vowel [ɑ] as a single vowel category because Japanese has only one counterpart, low mid vowel [a] (e.g., Nishi & Kewley-Port, 2007)
- **Vowels (Diphthongs):** Although English diphthongs are pronounced within a single syllable, Japanese listeners tend to substitute English diphthongs with long vowels ([boot] for [bout]) or mispronounce them with two distinguished syllables ([ka] + [u] for [kaʊ]) (Ohata, 2004).
- **Liquids:** Japanese listeners tend to perceive the English [l]-[r] contrast as the Japanese counterpart, alveolar tap [ɾ] (Iverson et al., 2003).

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

- **Fricatives:** Japanese has only three fricative sounds [s, z, h], resulting in difficulties for English-specific fricatives [f, v, θ, ð, s, z, ʃ, ʒ].
- **Affricates:** Japanese listeners have difficulties in perceiving [si] and [ti] because of the allophonic variations in the Japanese systems, /s/ → [ʃ] / i, /t/ → [tʃ] / i.

As described in Slevc and Miyake (2006), the test materials used in the current study featured most of the problematic English-specific phonological contrasts for Japanese listeners (for the test stimuli, see Supporting Information B, Supplementary Materials). The materials were read by a male native speaker of General American English and recorded at the sampling rate of 44.1 kHz with 16-bit quantization. Using headphones, participants listened to all the samples in a fixed order via MATLAB. For each trial, two alternatives were orthographically displayed on a computer screen, and participants were asked to choose which word they had heard.

Although the materials covered a list of the problematic English contrasts (summarized above), the number of presentations of each contrast was not equally distributed. The stimuli were played in a fixed order because we followed the procedure in Slevc and Miyake (2006). This could have resulted in an order effect. According to the results of corpus-based analyses, the target words comprised both frequent and infrequent words (1k to 15k frequency bands; Cobb, 2020). We did not further conduct any post-hoc analyses of the extent to which participants had familiarity with each lexical item. Note that the word frequency and familiarity of stimuli may have affected listeners' performance (Flege, Takagi, & Mann, 1996). Finally, the target stimuli were embedded in two different contexts ($n = 26$ in isolated words vs. $n = 25$ in sentences). To control for the phonological, lexical, and task status of each test item, a mixed-effects binomial logistic regression analysis was conducted. Participants' performance per item ($n = 51$ items per participant) was used as dependent variables with item ID (1-70) (and participant ID [1-68]) as random effects (see the Results section).

Results

Relationships between auditory processing and memory profiles

As summarized in Supporting Information C (Supplementary Materials), participants' auditory processing and memory abilities ranged widely. A normality test (Kolmogorov-Smirnov) found the composite auditory processing scores to show positive skewness ($D = .101$, $p = .074$), and so, following the procedure in the previous literature (e.g., Kachlicka et al., 2019), the raw scores were submitted to a square root transformation. The resulting scores did not significantly deviate from the normal distribution ($D = .097$, $p = .170$).

In order to examine the presence of any latent factors underlying a total of five perceptual-cognitive abilities (auditory processing, phonological short-term memory, complex working memory, declarative and procedural memory), exploratory factor analysis was conducted with Varimax rotation. We confirmed the factorability of the entire dataset via two tests: the Bartlett's test of sphericity ($\chi^2 = 26.068$, $p = .004$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (.608). There were only two factors with eigenvalues beyond 1.0. The model accounted for 56.27% of the total variance in the participants' abilities. In line with Hair, Anderson, Tatham, and Black's (1998) recommendations, the cut-off value for the practically significant factor loadings was set to 0.4.

Table 2. Summary of a Two-Factor Solution Based on a Factor Analysis of Auditory Processing and Memory Scores

	Factor 1: Explicit memory	Factor 2: Implicit audition
Cumulative%	34.40%	56.27%
Auditory processing[a]	−.071	**−.752**
Phonological short-term memory	**.627**	.271
Complex working memory	**.830**	.119
Declarative memory	**.695**	−.134
Procedural memory	.048	**.753**

Note. All loadings > .4 were highlighted in bold; [a] lower values indicate more precise auditory processing

As shown in Table 2, Factor 1 could be considered to index participants' explicit memory abilities: as the tasks (nonword repetition, operation span, paired associates) required them to remember, associate, and elaborate sensory information with awareness. Factor 2 could be interpreted as participants' relatively implicit perception as the tasks (auditory discrimination, serial reaction time) led them to implicitly detect and learn statistical distributions of numeral strings without awareness. Taken together, the findings suggest that whereas auditory discrimination performance can be considered as an independent construct compared to explicit memory abilities, it may overlap with implicit statistical learning abilities.

Auditory processing, memory, experience, and L2 speech proficiency

The next objective of the statistical analyses was to examine whether, to what degree, and how auditory processing alone can explain the variances in participants' L2 speech perception proficiency when their cognitive and experience profiles were statistically controlled for. According to the descriptive results of participants' identification scores, their L2 speech proficiency widely varied ($M = 71.0\%$, $SD = 7.6$; $Range = 49.0–88.2$). To control for the random effects of participants' performance per item, a mixed-effects binomial logistic regression analysis was performed using the lmer functions from the lme package (Version 1.1-21; Bates, Maechler, Bolker, & Walker, 2011) in the R statistical environment.

Participants' binary accuracy scores for each stimulus (0 for incorrect, 1 for correct) were used as dependent variables. Fixed effects included one auditory processing factor (composite auditory processing scores), four memory factors (phonological short-term memory, complex working memory, declarative memory, procedural memory), and three experience factors (length of foreign language education, pronunciation training, immersion experience). The two binary predictors (the presence/absence of pronunciation training and immersion experience) were treatment-coded. Variance Inflation Factor values among the six predictors did not yield multicollinearity problems (1.180-1.416). Random effects including participant ID (1-70) and stimulus ID (1-51) were entered.

The adequacy of the sample size was examined for the analysis involving cross-random factors of participants ($N = 70$) and stimuli ($N = 51$), resulting in a total of 3570 observations. In light of

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

Table 3. Summary of the Mixed-Effects Model Explaining the Perceptual, Cognitive, and Biographical Correlates of L2 Speech Perception Proficiency

Model 1	Fixed effects	Estimate (Standardized β)	*SE*	*t*	*p*
	Intercept	1.328	0.272	4.880	< .001
	Auditory processing[a]	−0.143	0.063	−2.236	.025*
	Phonological short-term memory	0.186	0.065	2.843	.004*
	Complex working memory	−0.066	0.069	−0.952	.341
	Declarative memory	0.035	0.063	0.555	.579
	Procedural memory	0.055	0.062	0.899	.368
	Length of foreign language education	0.042	0.064	0.648	.516
	Pronunciation training (yes, no)	0.056	0.128	0.439	.660
	Immersion experience (yes, no)	0.475	0.149	3.187	.001*
	Random effects	Variance	*SD*		
	Participant	0.104	0.323		
	Items	3.253	1.804		
	Conditional R^2	Marginal R^2			
	.148	.556			
Model 2	Fixed effects	Estimate (Standardized β)	*SE*	*t*	*p*
	Intercept	1.351	0.265	5.086	< .001*
	Auditory processing[a]	−0.158	0.065	−2.44	.014*
	Phonological short-term memory	0.124	0.069	1.8	.071†
	Complex working memory	−0.050	0.066	−0.751	.452
	Declarative memory	0.031	0.06	0.455	.649
	Procedural memory	0.019	0.060	0.315	.752
	Immersion experience (yes, no)	0.662	0.150	4.413	< .001*
	Auditory × immersion	0.107	0.156	0.682	.495
	Phonological short-term × immersion	0.325	0.187	1.734	.083†
	Complex working × immersion	−0.174	0.292	−0.598	.550
	Declarative × immersion	0.112	0.146	0.765	.444
	Procedural × immersion	0.399	0.202	1.979	.047*
	Random effects	Variance	*SD*		
	Participant	0.078	0.279		
	Items	3.252	1.803		
	Conditional R^2	Marginal R^2			
	.200	.513			

Note. * indicates $p < .05$; † indicates $p < .10$; [a]Lower values indicate more precise auditory processing

Westfall, Kenny and Judd's (2014) recommended procedure for calculating statistical power for a crossed design model, to reach a medium effect size of $d = 0.5$, the model resulted in relatively strong power, .990. The figure here substantially exceeds Larson-Hall's (2015) field-specific benchmark, .700.

As summarized in Table 3, Model 1 explained 14.8% of the variance in participants' L2 speech proficiency. Auditory processing, phonological short-term memory, and the immersion experience factor were identified as significant predictors. According to the standardized β values, the results indicated that participants' L2 speech proficiency was primarily determined by immersion experience ($\beta = .475$), followed by phonological short-term memory ($\beta = .186$) and auditory processing ($\beta = -.143$).

Given that the immersion experience factor (yes or no) was found to be the strongest predictor of L2 speech learning, there was a possibility that the perceptual and cognitive correlates of L2 speech learning may have differed between those with or without such experience. To further explore this, Model 2 was constructed to include both main and interaction terms. According to the results (see Table 3), whereas auditory processing remained significant ($\beta = -.158$, $p = .014$), the main and interaction effects of phonological short-term memory marginally reached significance ($\beta = .124, .187$, $p = .071, .083$). Interestingly, the interaction effects of procedural memory became significant ($\beta = .202$, $p = .047$). The results of post-hoc Spearman correlation analyses showed that procedural memory was significantly associated with experienced

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

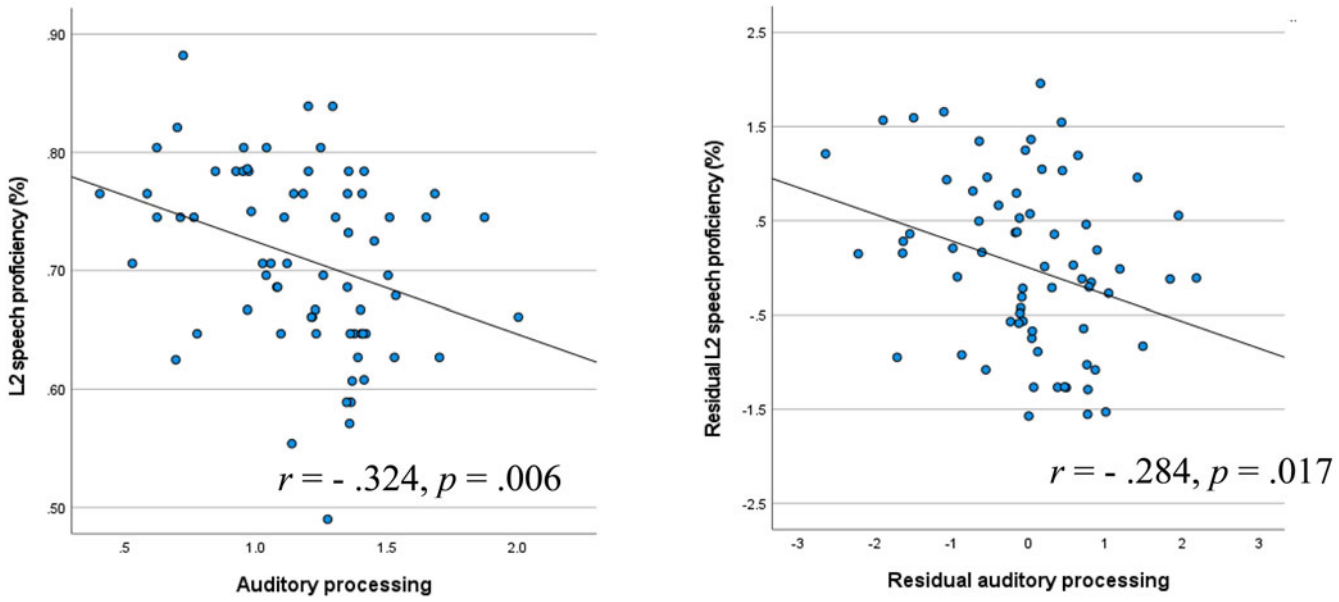


Fig. 1. Scatterplots displaying the simple correlation between L2 speech proficiency and auditory processing (Left) and the partial correlation between L2 speech proficiency and auditory processing when the variables related to experience (immersion) and cognitive abilities (phonological short-term memory, complex working memory, declarative memory, and procedural memory) controlled for (Right). Lower auditory processing scores indicate more precise auditory acuity.

participants' L2 speech proficiency ($r = .584$, $p = .009$), whereas the relationship was not significant among inexperienced participants ($r = .106$, $p = .459$).

In sum, the findings suggest (a) that auditory processing factors could make a UNIQUE contribution to the outcomes of L2 speech learning even after learners' individual differences in immersion experience and memory abilities were controlled for; and that (b) the predictive power of cognitive abilities (phonological short-term, procedural memory) tend to be clear especially when the analyses focus on students with immersion experience. The relationship between participants' auditory processing and proficiency scores (with and without the experience-related and cognitive factors) was visually summarized (see Figure 1).

Discussion

An emerging paradigm suggests that individual differences in auditory processing play a key role in determining the rate and ultimate attainment of post-pubertal L2 speech learning (Kachlicka et al., 2019; Saito & Tierney, forthcoming; Sun et al., 2021). Examining 70 college-level Japanese speakers of English with varied experience and proficiency levels, the current investigation explores whether auditory processing overlaps or differs from other cognitive abilities (working, declarative, and procedural memory); and the extent to which degree auditory processing can UNIQUELY explain the outcomes of L2 speech learning when participants' memory and experience profiles are factored out.

Constructs of auditory processing

The results of the factor analyses demonstrated that participants' auditory processing scores clustered with implicit statistical learning abilities (measured via serial reaction time) but were distinguishable from the other explicit memory abilities (measured via nonword repetition, operation span, and paired associates). The results here lend an empirical support to the view that auditory processing could be a multifaceted phenomenon – as the task design used to assess this ability (adaptive A× B discrimination with a fixed baseline) may tap into other cognitive abilities (Snowling, Gooch, McArthur & Hulme, 2018). Specifically, our results suggest that the auditory discrimination task may represent one's broader abilities to not only perceive subtle acoustic differences (i.e., auditory acuity) but also to detect and integrate the statistical distributions of target vs. anchor stimuli (i.e., implicit cognition; Ahissar et al., 2006).

In the discrimination task, one baseline stimulus was used as an anchor and played at every trial. As participants engaged in more opportunities to compare the acoustic differences and similarities between the anchor and target stimuli, they were induced to conduct a statistical learning of the acoustical properties of the prototype so that they could discriminate it from the target stimuli more accurately and promptly. In essence, our findings were in line with Raviv et al.'s hypothesis that auditory processing and implicit cross-trial statistical learning are interwoven to some degree (Raviv et al., 2012; Jaffe-Dax et al., 2017).

Another crucial point is that what was measured as auditory processing, via the discrimination task (auditory acuity, implicit cognition), was independent of participants' abilities to remember, analyze, and elaborate on sensory information WITH awareness (explicit memory). This could be arguably because the nature of the task (A × B discrimination) placed only limited demands on short-term memory capacity, given the relatively short length of each sample (500 ms). Thus, it is reasonable to say that the auditory perception ability measured via the discrimination task could be theoretically DISTINCT from explicit working and declarative memory (but somewhat overlapping with implicit procedural memory).

Auditory processing, memory, experience & L2 speech acquisition

Not surprisingly, the results of mixed-effects modeling analyses showed that Japanese university students' L2 outcomes could be

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

mainly determined by the presence of participants' study-abroad experience as reported in the existing literature (e.g., Muñoz, 2014). However, the remaining variance was associated with participants' perceptual-cognitive aptitude profiles, including auditory processing (Kachlicka et al., 2019), phonological short-term memory (Darcy et al., 2015), and procedural memory (Linck et al., 2013). Crucially, auditory processing remained as a significant predictor even after the experience and memory factors were partialled out. The findings here support the emerging paradigm that auditory processing UNIQUELY relates to acquisition regardless of participants' experience and cognitive states (Kachlicka et al., 2019). Given that domain-general auditory processing has been ignored in the existing language aptitude frameworks (e.g., Hi-Lab for Linck et al., 2013), future researchers are strongly recommended to include both auditory processing AND cognitive abilities in test batteries, in order to provide a more comprehensive picture of the aptitude effects in post-pubertal L2 speech learning (cf. Zheng, Saito & Tierney, 2021 for music aptitude vs. auditory processing).

It is noteworthy that the interaction effects of immersion experience were found significant for procedural memory and marginally significant for phonological short-term memory (but not for auditory processing). The results suggest that these cognitive abilities matter for L2 speech acquisition especially when learners engage in more conversational and communicatively authentic input in immersion settings, where they are encouraged to not only detect but also access/use auditory information with interlocutors on a regular basis (Faretta-Stutenberg & Morgan-Short, 2018). At the same time, however, these findings should be interpreted somewhat cautiously. There is a growing amount of evidence showing that auditory processing is a primary determinant of post-pubertal L2 speech learning in an interactive, meaningful and immersive setting, but not necessarily in a form-oriented, foreign language classroom context (Saito et al., 2020b; Saito et al., 2021a, 2021b). The topic (i.e., the perceptual and cognitive correlates of L2 speech acquisition) needs to be replicated especially among more advanced L2 learners with ample immersion experience (cf. Saito et al., 2020a).

Future directions

The current study derived two tentative conclusions: (a) that auditory perception skills were a relatively independent construct compared to other cognitive abilities (e.g., working, declarative, and procedural memory); and (b) that both perception and cognition skills serve as two SEPARATE constructs of aptitude relevant to post-pubertal L2 speech proficiency. In conjunction with the exploratory nature of the current project, there are a range of promising directions that future studies can further pursue to disentangle the complex relationship between auditory processing, cognitive abilities, experience, and L2 speech acquisition.

First, although auditory processing was measured only behaviourally in the current study, some scholars have proposed the use of electrophysiological measures to assess participants' neural encoding of synthesized sounds at preconscious levels, finding preliminary links to a range of L2 proficiency outcomes (e.g., Kachlicka et al., 2019; Saito et al., 2019, 2020a; Sun et al., 2021). It would be intriguing to investigate to what degree the neural and behavioural assessments of auditory processing DIFFERENTIALLY relate to a range of executive functions and L2 skills.

Second, following the theoretical and methodological paradigms in the existing literature (e.g., Flege & Bohn, 2021), a forced-choice identification task was adopted to index participants' L2 speech proficiency. However, although segmental perception is thought to serve as a cornerstone of acquisition, nevertheless L2 speech proficiency can be characterized as a multilayered phenomenon that needs to be examined from multiple angles (e.g., perception vs. production; segmentals vs. suprasegmentals; controlled vs. spontaneous; phoneme vs. word levels; see Saito & Plonsky, 2019 for a conceptual summary). Thus, future studies can further delve into how learners with diverse auditory and cognitive aptitude not only perceive but also produce L2 sounds: at various processing levels (controlled vs. spontaneous); at various time scales (segmental vs. suprasegmental); and in various contexts (frequent vs. infrequent words).

One limitation resulting from our use of a forced-choice perception task is that this method cannot assess the extent to which listeners perceive differences between speech sounds in a gradient versus categorical manner. Electrophysiology research suggests that, across listeners, continuous information about acoustic characteristics of speech sounds – such as voice onset time – is retained both at perceptual and post-perceptual stages (Toscano, McMurray, Dennhardt & Luck, 2010). However, there are individual differences between listeners in the degree to which they retain continuous information about speech, with some perceiving speech sound contrasts in a more categorical manner and others in a more gradient manner. These individual differences in gradient perception can be revealed with a Visual Analogue Scaling (VAS) task, in which participants are asked to rate the extent to which a sound resembles one of two speech sounds on a continuous scale: participants who only use the ends of the scale perceive speech more categorically, while participants who use the entire range more or less equally are more gradient perceivers (Kapnoula, Winn, Kong, Edwards & McMurray, 2017; Kapnoula & McMurray, 2021). More gradient perceivers make more flexible use of acoustic information when perceiving speech: they are more likely to integrate across multiple cues and are better able to recover from misleading information (Kapnoula, Edwards & McMurray, 2021). More gradient listeners, therefore, may be better able to formulate new perceptual strategies appropriate for their L2, despite having already developed alternate perceptual strategies tuned to their L1. This hypothesis could be tested in future research using the VAS task.

Third, the current study focused on the three experience-related variables – i.e., length of instruction, study-abroad, and pronunciation instruction. However, future studies can further survey precisely how often EFL learners are exposed to the target language on a daily basis per modality (listening, speaking, reading, and writing) and setting (social, home, and school). In order to design such multilayered research, scholars need to approach this topic from both quantitative and qualitative paradigms by using not only retrospective questionnaires but also daily learning log instruments (cf. Ranta & Meckelborg, 2013).

Fourth, most of the perceptual-cognitive abilities in this study could be considered domain-general (with the exception of phonological short-term memory). With a view towards a full-fledged understanding of the mechanisms underlying advanced L2 speech acquisition, future research can examine how domain-general aptitude is related to domain-SPECIFIC aptitude (e.g., phonemic coding; Saito et al., 2019), and how both domain-general and domain-specific measures can predict the incidence of high-level L2 proficiency (Campbell & Tyler, 2018). Fourth, it needs to be acknowledged that most of the literature (including the current investigation) is cross-sectional in nature. To prove the CAUSAL

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

relationship between auditory processing, cognitive abilities, experience, and L2 speech learning, a longitudinal study is strongly called for. One promising direction concerns aptitude-treatment interaction: future studies can investigate how those with different auditory and cognitive aptitude can differentially enhance both perception and production skills when they immerse in an L2 speaking environment for a prolonged period of time (e.g., Sun et al., 2021), and/or when they receive different types of instruction (e.g., Perrachione et al., 2011 for perception-based training [high variability phonetic training]; Shao, Saito & Tierney, 2022 for production-based training [shadowing]).

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728922000153

Acknowledgments. We gratefully acknowledge insightful comments from anonymous *Bilingualism: Language and Cognition* reviewers on earlier versions of the manuscript. This research was supported by two external grants from ESRC-AHRC UK-Japan SSH Connections Grant (ES/S013024/1; awarded to Saito, Tierney, Révész, Suzuki, Jeong, & Sugiura) and Leverhulme Research Grant (RPG-2019-039; awarded to Saito & Tierney), and two internal grants from the Graduate School of International Cultural Studies (awarded to Jeong), and the Institute of Development, Aging, and Cancer, Tohoku University (awarded to Saito, Sugiura, and Jeong).

References

Abrahamsson N and Hyltenstam K (2009) Age of Onset and Nativelikeness in a Second Language: Listener Perception Versus Linguistic Scrutiny. *Language Learning* **59**(2), 249–306. https://doi.org/10/fmfptp

Ahissar M, Lubin Y, Putter-Katz H and Banai K (2006) Dyslexia and the failure to form a perceptual anchor. *Nature Neuroscience* **9**(12), 1558–1564. https://doi.org/10/bwxgzw

Baddeley A (1993) Short-term phonological memory and long-term learning: A single case study. *European Journal of Cognitive Psychology* **5**(2), 129–148. https://doi.org/10.1080/09541449308520112

Baddeley A (2003) Working memory: Looking back and looking forward. *Nature Reviews Neuroscience* **4**(10), 829–839. https://doi.org/10.1038/nrn1201

Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, … & Grothendieck G (2011) Package 'lme4'. *Linear mixed-effects models using S4 classes. R package version* **1**(6).

Bidelman GM, Gandour JT and Krishnan A (2009) 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. https://doi.org/10/fnb3c3

Boets B, Vandermosten M, Poelmans H, Luts H, Wouters J and Ghesquière P (2011) Preschool impairments in auditory processing and speech perception uniquely predict future reading problems. *Research in Developmental Disabilities* **32**(2), 560–570. https://doi.org/10/fg7rq9

Boets B, Wouters J, van Wieringen A, De Smedt B and Ghesquière P (2008) Modelling relations between sensory processing, speech perception, orthographic and phonological ability, and literacy achievement. *Brain and Language* **106**(1), 29–40. https://doi.org/10/c6qsf5

Campbell KL and Tyler LK (2018) Language-related domain-specific and domain-general systems in the human brain. *Current Opinion in Behavioral Sciences* **21**, 132–137. https://doi.org/10/ggqgcp

Casini L, Pech-Georgel C and Ziegler JC (2018) It's about time: Revisiting temporal processing deficits in dyslexia. *Developmental Science* **21**(2), e12530. https://doi.org/10/f9sqs6

Cobb T (2020) *The compleat lexical tutor*. Retrieved from http://www.lextutor.ca/vp/

Conway A, Jarrold C, Kane M, Miyake A and Towse J (eds.) (2007) *Variation in working memory*. Oxford University Press.

Corriveau KH and Goswami U (2009) Rhythmic motor entrainment in children with speech and language impairments: Tapping to the beat. *Cortex* **45**(1), 119–130. https://doi.org/10/cvsxnq

Cumming R, Wilson A and Goswami U (2015) Basic auditory processing and sensitivity to prosodic structure in children with specific language impairments: A new look at a perceptual hypothesis. *Frontiers in Psychology* **6**. https://doi.org/10/gg6xtj

Cutler A and Butterfield S (1992) Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language* **31**(2), 218–236. https://doi.org/10/cg8f4f

Darcy I, Park H and Yang C-L (2015) Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences* **40**, 63–72. https://doi.org/10/f7hbt9

Derwing TM and Munro MJ (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/gbfknb

Doughty CJ (2019) Cognitive language aptitude. *Language Learning* **69**(S1), 101–126. https://doi.org/10/ggcp55

Faretta-Stutenberg M and Morgan-Short K (2018) The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research* **34**(1), 67–101. https://doi.org/10/dswc

Flege J and Bohn O-S. (2021) The revised Speech Learning Model. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress*. Cambridge: Cambridge University Press.

Flege JE and Liu S (2001) The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* **23**(4), 527–552. https://doi.org/10/gg6xtv

Flege JE, Takagi N and Mann V (1996) Lexical familiarity and English-language experience affect Japanese adults' perception of /r/ and /l/. *Journal of Acoustical Society of America* **99**, 1161–1173.

Gathercole SE, Frankish CR, Pickering SJ and Peaker S (1999) Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **25**(1), 84–95. https://doi.org/10.1037/0278-7393.25.1.84

Gibson LY, Hogben JH and Fletcher J (2006) Visual and auditory processing and component reading skills in developmental dyslexia. *Cognitive Neuropsychology* **23**(4), 621–642. https://doi.org/10/d9g86j

Gordon RL, Shivers CM, Wieland EA, Kotz SA, Yoder PJ and McAuley JD (2015) Musical rhythm discrimination explains individual differences in grammar skills in children. *Developmental Science* **18**(4), 635–644. https://doi.org/10/f7dzjn

Goswami U (2015) Sensory theories of developmental dyslexia: Three challenges for research. *Nature Reviews Neuroscience* **16**(1), 43–54. https://doi.org/10/gf9qkc

Goswami U, Wang H-LS, Cruz A, Fosker T, Mead N and Huss M (2010) Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *Journal of Cognitive Neuroscience* **23**(2), 325–337. https://doi.org/10/ct2mgm

Haake C, Kob M, Willmes K and Domahs F (2013) Word stress processing in specific language impairment: Auditory or representational deficits? *Clinical Linguistics & Phonetics* **27**(8), 594–615. https://doi.org/10/gg6xbx

Hair JF, Anderson RE, Tatham RL and Black WC (1998) *Multivariate data analysis* (5th ed.) Upper Seddle River, New Jersey, USA: Prentice-Hall International, Inc.

Hamrick P, Lum JAG and Ullman MT (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://doi.org/10/gc3bz2

Hornickel J and Kraus N (2013) Unstable representation of sound: A biological marker of dyslexia. *Journal of Neuroscience* **33**(8), 3500–3504. https://doi.org/10/f4ph3w

Hu X, Ackermann H, Martin JA, Erb M, Winkler S and Reiterer SM (2013) Language aptitude for pronunciation in advanced second language (L2) Learners: Behavioural predictors and neural substrates. *Brain and Language* **127**(3), 366–376. https://doi.org/10.1016/j.bandl.2012.11.006

Iverson P, Kuhl PK, Akahane-Yamada R, Diesch E, Tohkura Y, Kettermann A and Siebert C (2003) A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition* **87**(1), B47–B57. https://doi.org/10/c66pbb

Jaffe-Dax S, Frenkel O and Ahissar M (2017) Dyslexics' faster decay of implicit memory for sounds and words is manifested in their shorter neural adaptation. *ELife* **6**, e20557. https://doi.org/10/f9mpdb

Jasmin K, Sun H and Tierney AT (2021) Effects of language experience on domain-general perceptual strategies. *Cognition* **206**, 104481. https://doi.org/10.1101/2020.01.02.892943

Joanisse MF and Seidenberg MS (1998) Specific language impairment: A deficit in grammar or processing? *Trends in Cognitive Sciences* **2**(7), 240–247. https://doi.org/10/bhvtpc

Jost LB, Eberhard-Moscicka AK, Pleisch G, Heusser V, Brandeis D, Zevin JD and Maurer U (2015) Native and non-native speech sound processing and the neural mismatch responses: A longitudinal study on classroom-based foreign language learning. *Neuropsychologia* **72**, 94–104. https://doi.org/10/f7dmtx

Kachlicka M, Saito K and 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/gg6xtx

Kapnoula E and McMurray B (2021) Idiosyncratic use of bottom-up and top-down information leads to differences in speech perception flexibility: converging evidence from ERPs and eye-tracking. *Brain and Language* **223**, 105031.

Kapnoula EC, Winn MB, Kong EJ, Edwards J and McMurray B (2017) Evaluating the sources and functions of gradiency in phoneme categorization: An individual differences approach. *Journal of Experimental Psychology: Human Perception and Performance* **43**(9), 1594.

Kapnoula EC, Edwards J and McMurray B (2021) Gradient activation of speech categories facilitates listeners' recovery from lexical garden paths, but not perception of speech-in-noise. *Journal of Experimental Psychology: Human Perception and Performance* **47**, 578.

Kaufman SB, DeYoung CG, Gray JR, Jiménez L, Brown J and Mackintosh N (2010) Implicit learning as an ability. *Cognition* **116**(3), 321–340. https://doi.org/10/dbrzzv

Kempe V, Bublitz D and Brooks PJ (2015) Musical ability and non-native speech-sound processing are linked through sensitivity to pitch and spectral information. *British Journal of Psychology* **106**(2), 349–366. https://doi.org/10/gg6xbt

Knowlton BJ, Mangels JA and Squire LR (1996) A neostriatal habit learning system in humans. *Science* **273**(5280), 1399–1402. https://doi.org/10/c7g8mm

Kraus N and Banai K (2007) Auditory-processing malleability: Focus on language and music. *Current Directions in Psychological Science* **16**(2), 105–110. https://doi.org/10/b7hqwm

Kuhl PK (2004) Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience* **5**(11), 831–843. https://doi.org/10/dzjxck

Larson-Hall J (2015) A guide to doing statistics in second language research using SPSS and R. Routledge. https://doi.org/10.4324/9781315775661

Lengeris A and 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://doi.org/10/b2m4vj

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/bchrz4

Li S (2016) The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition* **38**(4), 801–842. https://doi.org/10.1017/S027226311500042X

Linck JA, Hughes MM, Campbell SG, Silbert NH, Tare M, Jackson SR, Smith BK, Bunting MF and Doughty CJ (2013) Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning* **63**(3), 530–566. https://doi.org/10/f46zgn

Linck JA, Osthus P, Koeth JT and Bunting MF (2014) Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin and Review* **21**(4), 861–883. https://doi.org/10/f6cmbz

Marslen-Wilson WD, Tyler LK, Warren P, Grenier P and Lee CS (1992) Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology Section A* **45**(1), 73–87. https://doi.org/10/bf2d86

McAllister R, Flege JE and Piske T (2002) The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of phonetics* **30**(2), 229–258. https://doi.org/10.1006/jpho.2002.0174

McArthur GM and Bishop DVM (2004) Which people with specific language impairment have auditory processing deficits? *Cognitive Neuropsychology* **21**(1), 79–94. https://doi.org/10/bj757c

McArthur GM and Bishop DVM (2005) Speech and non-speech processing in people with specific language impairment: A behavioural and electrophysiological study. *Brain and Language* **94**(3), 260–273. https://doi.org/10/d8krbd

Meara P (2005) LLAMA language aptitude tests: The manual. *Swansea: Lognostics*.

Mora JC and Valls-Ferrer M (2012) Oral fluency, accuracy, and complexity in formal instruction and study Abroad learning contexts. *TESOL Quarterly* **46**(4), 610–641. https://doi.org/10/gg6xtq

Morgan-Short K, Faretta-Stutenberg M, Brill-Schuetz KA, Carpenter H and Wong PCM (2014) Declarative and procedural memory as individual differences in second language acquisition. *Bilingualism* **17**(1), 56–72. https://doi.org/10/f5vt55

Moritz C, Yampolsky S, Papadelis G, Thomson J and Wolf M (2013) Links between early rhythm skills, musical training, and phonological awareness. *Reading and Writing* **26**(5), 739–769. https://doi.org/10/f4z7dq

Mueller JL, Friederici AD and Männel C (2012) Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences* **109**(39), 15953–15958. https://doi.org/10.1073/pnas.1204319109

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/f6kpmc

Nishi K and Kewley-Port D (2007) Training Japanese listeners to perceive American English vowels: Influence of training sets. https://doi.org/10.1044/1092-4388(2007/103)

Nishino T and Watanabe M (2008) Communication-oriented policies versus classroom realities in Japan. *TESOL Quarterly* **42**(1), 133–138.

Norris D and McQueen JM (2008) Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review* **115**(2), 357–395. https://doi.org/10.1037/0033-295X.115.2.357

Ohata K (2004) Phonological differences between Japanese and English: Several potentially problematic areas of pronunciation for Japanese ESL/EFL learners. *Asian EFL Journal* [On–line], *6*(4). Available http://www.asian-efl-journal.com/december_2004_KO.php

Omote A, Jasmin K and Tierney A (2017) Successful non-native speech perception is linked to frequency following response phase consistency. *Cortex* **93**, 146–154. https://doi.org/10/gbskjf

Papagno C, Valentine T and Baddeley A (1991) Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language* **30**(3), 331–347. https://doi.org/10.1016/0749-596X(91)90040-Q

Perrachione TK, Lee J, Ha LYY and Wong PCM (2011) Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *The Journal of the Acoustical Society of America* **130**(1), 461–472. https://doi.org/10/c4mhmt

Peynircioglu ZF, Durgunoglu AY and Oney-Kusefoglu B (2002) Phonological awareness and musical aptitude. *Journal of Research in Reading* **25**, 68–80. doi:10.1111/1467-9817.00159.

Plonsky L and Oswald FL (2014) How Big Is "Big"? Interpreting Effect Sizes in L2 Research. *Language Learning* **64**(4), 878–912. https://doi.org/10.1111/lang.12079

Ranta L and Meckelborg A (2013) How much exposure to English do international graduate students really get? Measuring language use in a naturalistic setting. *The Canadian Modern Language Review* **69**, 1–33.

Raviv O, Ahissar M and Loewenstein Y (2012) How recent history affects perception: The normative approach and its heuristic approximation. *PLoS Comput Biol* **8**(10), e1002731. https://doi.org/10.1371/journal.pcbi.1002731

Reiterer SM, Hu X, Erb M, Rota G, Nardo D, Grodd W, Winkler S and Ackermann H (2011) Individual Differences in Audio-Vocal Speech Imitation Aptitude in Late Bilinguals: Functional Neuro-Imaging and Brain Morphology. *Frontiers in Psychology* **2**. https://doi.org/10.3389/fpsyg.2011.00271

Révész A (2012) Working memory and the observed effectiveness of recasts on different L2 outcome measures. *Language Learning* **62**(1), 93–132. https://doi.org/10/fzbsdz

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press

Rosen S (2003) Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything? *Journal of Phonetics* **31**(3), 509–527. https://doi.org/10/bs4nfs

Russo NM, Skoe E, Trommer B, Nicol T, Zecker S, Bradlow A and Kraus N (2008) Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. *Clinical Neurophysiology* **119**(8), 1720–1731. https://doi.org/10/drpcpw

Saito K (2014) Experienced teachers' perspectives on priorities for improved intelligible pronunciation: The case of Japanese learners of English. *International Journal of Applied Linguistics* **24**, 250–277.

Saito K (2018) Advanced segmental and suprasegmental acquisition. In Malovrh P and Benati A (Eds.). *The handbook of advanced proficiency in second language acquisition*. Hoboken, NJ: Wiley Blackwell, pp. 282–303.

Saito K and Hanzawa K (2016) Developing second language speaking proficiency in foreign language classrooms: The role of the length and focus of instruction and individual differences. *Applied Psycholinguistics* **37**, 813–840. https://doi.org/10.1017/S0142716415000259

Saito K, Sun H and Tierney A (2019) Explicit and implicit aptitude effects on second language speech learning: Scrutinizing segmental and suprasegmental sensitivity and performance via behavioural and neurophysiological measures. *Bilingualism: Language and Cognition* **22**(5), 1123–1140. https://doi.org/10.1017/S1366728918000895

Saito K, Kachlicka M, Sun H and Tierney A (2020a) Domain-general auditory processing, age, experience, and post-pubertal L2 speech learning: A behavioral and neurophysiological investigation. *Journal of Memory and Language* **115**, 104168. https://doi.org/10.1016/j.jml.2020.104168

Saito K, Sun H and Tierney A (2020b) Domain-general auditory processing as a perceptual-cognitive anchor of L2 pronunciation learning in adulthood: A longitudinal study. *Applied Psycholinguistics* **41**, 1083–1112. https://doi.org/10.1017/S0142716420000491

Saito K, Sun H and Tierney A (2020c) *Brief report: Test-retest reliability of explicit auditory processing measures*. https://doi.org/10.1101/2020.06.12.149484

Saito K, Suzukida Y, Tran M and Tierney A (2021a) Domain-general auditory processing partially explains L2 speech learning in classroom settings: A review and generalization study. *Language Learning* **71**, 669–715. https://doi.org/10.1111/lang.12447

Saito K, Sun H, Kachlicka M, Robert J, Nakata N and Tierney A (2021b) Domain-general auditory processing explains multiple dimensions of L2 acquisition in adulthood. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263120000467

Saito K, Macmillan K, Kroeger S, Magne V, Takizawa K, Kachlicka M and Tierney A (2022) Roles of domain-general auditory processing in spoken second language vocabulary attainment in adulthood. *Applied Psycholinguistics*.

Saito K and 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 K and Tierney A (forthcoming) Domain-general auditory processing as a conceptual and measurement framework for second language speech learning aptitude: Systematic review and test development.

Shao Y, Saito K and Tierney A (2022) How does having a good ear promote instructed second language pronunciation development? Roles of domain-general auditory processing in choral repetition training. *TESOL Quarterly*. https://doi.org/10.1002/tesq.3120

Sharma M, Purdy SC and Kelly AS (2009) Comorbidity of auditory processing, language, and reading disorders. *Journal of speech, language, and hearing research : JSLHR* **52**(3), 706–722. https://doi.org/10.1044/1092-4388(2008/07-0226)

Silbert NH, Smith BK, Jackson SR, Campbell SG, Hughes MM and Tare M (2015) Non-native phonemic discrimination, phonological short term memory, and word learning. *Journal of Phonetics* **50**, 99–119. https://doi.org/10.1016/j.wocn.2015.03.001

Skoe E, Krizman J, Anderson S and Kraus N (2015) Stability and plasticity of auditory brainstem function across the lifespan. *Cerebral Cortex* **25**(6), 1415–1426. https://doi.org/10/f7dq52

Slevc LR and Miyake A (2006) Individual Differences in Second-Language Proficiency: Does Musical Ability Matter? *Psychological Science* **17**(8), 675–681. https://doi.org/10.1111/j.1467-9280.2006.01765.x

Smith JO (2007) *Introduction to digital filters: With audio applications*. Julius Smith.

Snowling MJ, Gooch D, McArthur G and Hulme C (2018) Language skills, but not frequency discrimination, predict reading skills in children at risk of dyslexia. *Psychological Science*. https://doi.org/10/gdnd86

Sun H, Saito K and 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. https://doi.org/10.1017/S0272263120000649

Tallal P (2004) Improving language and literacy is a matter of time. *Nature Reviews Neuroscience* **5**(9), 721–728. https://doi.org/10/fb93cg

Tierney A and Kraus N (2014) Auditory-motor entrainment and phonological skills: Precise auditory timing hypothesis (PATH). *Frontiers in Human Neuroscience* **8**. https://doi.org/10/gg6xtk

Tierney A, White-Schwoch T, MacLean J and Kraus N (2017) Individual differences in rhythm skills: Links with neural consistency and linguistic ability. *Journal of cognitive neuroscience* **29**, 855–868. https://doi.org/10.1162/jocn_a_01092

Toscano JC, McMurray B, Dennhardt J and Luck SJ (2010) Continuous perception and graded categorization: Electrophysiological evidence for a linear relationship between the acoustic signal and perceptual encoding of speech. *Psychological science* **21**(10), 1532–1540.

Tulving E (1993) What Is Episodic Memory? *Current Directions in Psychological Science* **2**(3), 67–70. https://doi.org/10/fg4d84

Turner ML and Engle RW (1989) Is working memory capacity task dependent? *Journal of Memory and Language* **28**(2), 127–154. https://doi.org/10/ddtm9j

Ullman MT (2004) Contributions of memory circuits to language: The declarative/procedural model. *Cognition* **92**(1–2), 231–270. https://doi.org/10/b25czn

Ullman MT and Lovelett JT (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/gg6xbk

Unsworth N, Heitz RP, Schrock JC and Engle RW (2005) An automated version of the operation span task. *Behavior Research Methods* **37**(3), 498–505. https://doi.org/10/cgkprg

Westfall J, Kenny DA and Judd CM (2014) Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General* **143**, 2020–2045.

White-Schwoch T, Carr KW, Thompson EC, Anderson S, Nicol T, Bradlow AR, Zecker SG and Kraus N (2015) Auditory processing in noise: A Preschool biomarker for literacy. *PLOS Biology* **13**(7), e1002196. https://doi.org/10/f7qb8f

Won JH, Tremblay K, Clinard CG, Wright RA, Sagi E and Svirsky M (2016) The neural encoding of formant frequencies contributing to vowel identification in normal-hearing listeners. *The Journal of the Acoustical Society of America* **139**(1), 1–11. https://doi.org/10/f8qf9k

Wong PCM and Perrachione TK (2007) Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics* **28**(4), 565–585. https://doi.org/10/d9p5nv

Wright BA, Bowen RW and Zecker SG (2000) Nonlinguistic perceptual deficits associated with reading and language disorders. *Current Opinion in Neurobiology* **10**(4), 482–486. https://doi.org/10/d9jw4r

Yamaguchi A and Shimizu M (2011) *The relationship between phonological working memory and English proficiency of Japanese university EFL learners ii*.

Zhang Y-X, Moore DR, Guiraud J, Molloy K, Yan T-T and Amitay S (2016) Auditory Discrimination Learning: Role of Working Memory. *PLOS ONE* **11**(1), e0147320. https://doi.org/10/f8pqdj

Zheng C, Saito K and Tierney A (2021) Successful second language pronunciation learning is linked to precise auditory processing rather than music aptitude. *Second Language Research*. https://doi.org/10.1177/0267658320978493

https://doi.org/10.1017/S1366728922000153 Published online by Cambridge University Press