The Pedagogical Potential of Task-Based Language Teaching for Second Language Pronunciation: Roles of Individual Differences

by Jiying Xu

British Council's Master's Dissertation Awards 2024

Commendation

The Pedagogical Potential of Task-Based Language Teaching for Second Language Pronunciation: Roles of Individual Differences

Abstract

This research aimed to assess the effectiveness of task-based language teaching (TBLT) in enhancing the acquisition of second language speech skills, specifically in the context of Chinese learners' perception of English vowel pairs (/e/-/æ/ and /ɪ/-/iː/). Additionally, the study explored whether variations in auditory processing and working memory among participants influenced the impact of task-based instruction. The research involved 70 young adult English as a foreign language (EFL) learners, randomly divided into experimental and control groups. Both groups received a 30-minute task-based English session, and pre- and post-tests were administered to measure improvements in the perception of /e/-/æ/ and /ɪ/-/iː/ vowel pairs. Participants' language aptitude, including auditory processing and working memory, was assessed using three psychoacoustic discrimination tests and two visual and text-entry digit span tests, respectively. The results revealed two key findings: (a) the experimental group demonstrated a significant overall improvement of 9.05%, with specific gains of 8.6% for trained words and 7.8% for untrained words, accompanied by medium-to-large effect sizes ($\eta p^2 = 0.187$), and (b) variations in auditory processing, particularly in formant discrimination ability, and working memory partially accounted for the effectiveness of TBLT in enhancing second language segmental perception.

Keywords: *task-based instruction, language aptitude, auditory processing, working memory, second language speech teaching and learning, segmental perception*

Table of Contents

List of Tables

List of Figures

1. Introduction

Over the past few decades, research in second language acquisition (SLA) has extensively explored the benefits of task-based language teaching and learning (TBLT). This pedagogical approach has demonstrated its capacity to create conditions that engage cognitive processes crucial for second or foreign language (L2) acquisition (Robinson, 2011; Skehan, 2014). Notably, TBLT has been successful in directing learners' attention to desired language forms within meaningful communication, promoting the automaticity of language structures, and encouraging the use of more accurate, diverse, and complex language forms. However, most research on TBLT's efficacy has concentrated primarily on tasks related to grammar, vocabulary, and occasionally, pragmatic aspects (Solon et al., 2017).

Despite the wealth of research on the role of tasks and their features in fostering L2 development, there exists a significant gap in our understanding when it comes to their impact on the improvement of L2 pronunciation. Pronunciation is a fundamental component of L2 acquisition and effective communication skills, yet it has received relatively limited attention within the TBLT framework. While TBLT acknowledges the importance of pronunciation (Ellis, 2009) and some studies have recognized the occurrence of form-focused episodes (FFEs) related to pronunciation during task-based interactions (Loewen, 2005; Gurzynski-Weiss & Baralt, 2014), only a handful of studies have specifically investigated the connection between TBLT and pronunciation. Furthermore, the few studies that have touched upon pronunciation aspects (Ellis et al., 2001; Loewen, 2005) have generally treated pronunciation as an incidental aspect, often alongside grammatical and lexical considerations.

This dissertation sets out to address this research gap by exploring the effectiveness of TBLT in the realm of L2 speech teaching and learning, with a specific focus on English vowel perception. Vowel perception is a crucial aspect of L2 pronunciation, and its enhancement holds significant implications for achieving a higher level of L2 proficiency. Thus, one of the main aims of this study aims is to investigate the relationship between task-based activities and improvements in English vowel perception.

To contextualize this investigation further, it is important to acknowledge the substantial variability observed in adult L2 acquisition outcomes. Even when individuals engage in similar amounts and types of language practice, they frequently exhibit varying levels of L2 proficiency. This variability extends beyond mere exposure to the L2 and is influenced by individual perceptual and cognitive factors associated with explicit language learning, commonly referred to as language aptitude (Skehan, 2019; Li, 2016).

The ongoing debate within cognitive psychology, exploring the specificity of neural mechanisms involved in native language (L1) acquisition as compared to their applicability in various learning contexts (Campbell & Tyler, 2018 vs. Hamrick, Lum, & Ullman, 2018), provides a valuable context for comprehending the cognitive aspects of second language (L2) acquisition. One critical cognitive dimension within this debate is auditory processing, encompassing processes like encoding, representation, and retention of temporal and frequency characteristics of sounds. While prior research has established connections between auditory processing and short-term language skill acquisition (Kachlicka et al., 2019; Saito et al., 2020), it is equally imperative to investigate the relationship between explicit auditory processing skills, such as formant and pitch discrimination, and their impact on L2 speech perception and production.

Another vital cognitive facet engaged during various language-related activities, including L2 comprehension, production, and learning, is working memory. In the field of second language acquisition (SLA), particularly within the context of interactions, extensive research has focused on the role of working memory (WM) in the process of recognizing feedback and facilitating learning through interaction in task-based learning environments.

Furthermore, learner factors such as aptitude and working memory hold significant positions within Robinson's pedagogic task classification framework (Robinson, 2001a, 2001b, 2005, 2007a). Therefore, it is paramount to investigate how working memory can mediate the learning outcomes of task-based instruction. This exploration will shed light on the intricate interplay between cognitive functions, instructional methods, and language acquisition processes in the context of L2 learning.

In summary, this dissertation aims to bridge the gap in the existing literature by examining the effectiveness of task-based activities on English vowel perception within the context of language aptitude. The variability in L2 acquisition outcomes and the role of cognitive factors, particularly auditory processing and working memory, will be central to this investigation.

2. Literature Review

2.1. L2 Speech Teaching and Learning

In recent years, there has been a pronounced surge in scholarly interest directed towards the domain of second language (L2) speech acquisition and pedagogy. This heightened scholarly attention has led to an exploration of specific methodologies and approaches within this field. A synthesis study by Saito (2012) found that instruction effectively enhances segmental and suprasegmental aspects of L2 speech, with the type of instruction being an important variable impacting pronunciation performance at different processing levels. Over the past two decades, L2 pronunciation research has shifted its focus towards explicit instruction, defined as instruction distinct from decontextualized, nativism-focused language teaching (Saito & Plonsky, 2019).

Empirical studies consistently demonstrate the efficacy of explicit pronunciation instruction in improving intelligible and comprehensible L2 speech (Derwing & Munro, 2009; Thomson & Derwing, 2015; Saito & Plonsky, 2019; Saito et al., 2022). Sakai and Moorman's meta-analysis (2018) on 18 perception training studies revealed medium-sized improvements in perception ($d = 0.93$, $SD = 0.72$) and production abilities ($d = 0.89$, $SD = 0.61$) with explicit instruction. Saito's summary (2022) of five key high variability phonetic training (HVPT) studies on L2 English vowel acquisition demonstrates a noteworthy enhancement in participants' performance, indicating moderate progress in English /r/ and /l/ sounds (5–15% gains) and a moderate-to-large improvement in L2 English vowel acquisition (15–20% gains) following an average of 6 hours of explicit instruction.

Despite the effectiveness of explicit instruction, other methods like focus-on-form instruction (FFI) and recasts have also proven effective in L2 speech teaching. In Saito's study (Saito & Lyster, 2012), a 4-hour FFI led to medium-to-large improvements in perception (5.9% gain). Research has also explored the applicability of recasts to L2 phonological learning, consistently showing that pronunciation-focused corrective feedback (CF) enhances both segmental and suprasegmental accuracy (Saito, 2021).

As for the target of L2 speech teaching and learning, while some contend that L2 learners can achieve accent-free speech (Flege et al. 1995), a prevailing perspective among L2 speech researchers emphasizes the importance of setting more pragmatic goals, namely the acquisition of intelligible and comprehensible pronunciation (Munro & Derwing, 1995; Levis, 2005; Isaacs et al., 2018). There has been an increasing focus on the exploration of which aspects of pronunciation hold greater significance in achieving effective comprehensibility, regardless of the presence of an accent, as noted by Levis in 2005. This endeavor proves most effective when it directs its attention towards fundamental elements, encompassing not only individual sound components such as vowels and consonants but also overarching aspects of speech like stress, rhythm, and intonation. Research by Field in 2005, Hahn in 2004, and Munro & Derwing in 2006 has underscored the pivotal role of these segmental and suprasegmental features in substantially enhancing overall intelligibility.

Recent research in second language (L2) pronunciation has started to indicate that some pronunciation elements hold more significance in terms of how understandable the speech is perceived to be. It is advisable to prioritize teaching and learning those pronunciation features that have an impact on listeners' comprehension in the context of L2 speech instruction (Saito & Plonsky, 2019). These features encompass not only suprasegmental aspects such as lexical and sentence stress (Field, 2005; Hahn, 2004) but also segmental features, specifically those with high functional load (FL) (Munro & Derwing, 2006; Suzukida & Saito, 2021). FL, or Functional Load, represents an inventory of phonemic distinctions meticulously assessed and stratified in terms of their significance for effective communication. These differences have been carefully identified by closely examining pairs of words that have very similar sounds but differ by just one sound. These word pairs are commonly used in the language. Additionally, the analysis considered how certain sounds are pronounced differently in various English dialects spoken in different regions. It also looked at where in words these sound differences occur. In simpler terms, the statement highlights that researchers have thoroughly studied these specific language differences.

Employing this theoretical framework, the researchers proceeded to categorize these phonemic distinctions into two discrete tiers of functional significance: those rated from 10 to 6 were categorized as high Functional Load (high FL), whereas those rated from 5 to 1 were designated as low Functional Load (low FL) (Suzukida & Saito, 2021). Essentially, Munro and Derwing's research (2006) validated the notion that there exists diversity in the significance of acquiring specific speech sounds, including both highly distinctive and less distinctive segmental elements. This implies that educators ought to give precedence to mastering the highly distinctive segmental components (Munro et al., 2015). In the context of the Lingua Franca Core (LFC) framework, it was proposed that mastering core segmental elements is crucial for generating comprehensible spoken language (Jenkins, 2000). Previous studies have suggested that, while achieving native-like pronunciation in phonology remains challenging, adult learners can acquire specific consonants and vowels (Best & Tyler, 2007).

However, the outcomes vary for different speech sounds and are influenced by factors like individual learner characteristics, language pairings, and linguistic factors such as word frequency (Munro & Derwing, 2008). Nevertheless, it's worth noting that accuracy in perceiving and producing speech sounds plays a significant role in learners' comprehensibility and perceived accent, with comprehensibility being associated with all linguistic aspects and accent perception strongly linked to pronunciation, particularly segmental features, as opposed to vocabulary and grammar (Saito et al., 2016).

Furthermore, numerous research works have emphasized that variations in the way native and nonnative speakers perceive individual speech sounds can hinder the comprehension of a second language and potentially lead to challenges in recognizing words (Munro & Derwing, 2008). To successfully articulate distinct sounds unique to a second language, a speaker must initially possess the capacity to identify unfamiliar sounds within that language and establish a corresponding mental framework, known as a phonetic category (Lee et al., 2020).Some researchers argue that phonetic and phonological representations may originate from perceptual experiences (Flege, 1995), while others argue for an articulatory basis (Lee et al., 2020), it's undeniable that perception and production are closely interrelated and are both key to L2 speech learning. According to Flege et al. (1997), to articulate an L2-specific sound, learners must initially detect novel sounds within the L2 and form corresponding mental representations (phonetic categories). Thus, according to the central tenet that L2 speech acquisition is perception-based, enhancing learners' perception abilities may maximize the efficacy and outcomes of the acquisition process (Lee et al., 2020). As perception is recognized as a prerequisite to production, this paper will predominantly emphasize the role of instruction in improving learners' segmental perception by teaching two minimal pairs with high functional load which can be key to Chinese EFL learners communication process.

While explicit instruction and the target of instruction is of great importance to learners' L2 speech development, some findings also substantiate the notion that FFI achieves heightened efficacy when woven into communicative contexts (Celce-Murcia et al., 2010). To better understand the L2 speech teaching, studies in the past decades have also underscored the potential of combining form-focused instruction (FFI) with meaningful communication (Celce-Murcia et al., 2010; Saito, 2013; Ruan & Saito, 2023). FFI, defined as pedagogical efforts to draw learners' attention to language form, proves most effective in content-based and communicative language classrooms, where conveying meaningful messages takes precedence (Spada, 2011). The integration of form-focused instruction (FFI) within meaningful communicative contexts holds a pivotal position in task-based language teaching frameworks (Ellis, 2003; Long, 1985, 1998, 2015; Long & Norris, 2000; Robinson, 2011; Skehan, 1998). Recent research by Gordon (2021) suggests that learners exposed to FFI supplemented by communicative activities exhibited a significant enhancement in comprehensibility during the post-test phase. This noteworthy difference can be attributed to the integration of FFI with communicative elements, suggesting that the introduction of genuine communicative pressure, akin to real-life interaction, prompted more pronounced improvements in pronunciation.

Thus, within the realm of TBLT, a clear opportunity emerges to effect substantial change in L2 speech learning. By incorporating a communicative component that compels learners to allocate their cognitive resources towards meaning-focused tasks while concurrently fostering automaticity in pronunciation, TBLT promises a novel and promising avenue for advancing pronunciation instruction (Saito & Plonsky, 2019).

2.2. TBLT and L2 Speech

Task-Based Language Teaching (TBLT), or task-based instruction (TBI), is an approach that emphasizes meaningful task completion using authentic language input (Mora-Plaza et al., 2018). TBLT falls under communicative methodology, aiming to foster L2 learners' communicative competence through meaningful interactions when task-oriented (Ellis & Shintani, 2014). In TBLT, tasks serve as 'workplans' primarily focused on meaning, requiring learners to rely on their linguistic resources. These tasks typically involve an 'information gap' and a clear communicative outcome (Ellis, 2003). From an acquisitional perspective, TBLT is considered effective as it encourages L2 interaction and meaning negotiation, leading to the recognition of significant linguistic forms in input (Doughty & Williams, 1998; Schmidt, 2001). It also prompts more language-related episodes, where learners discuss specific linguistic features (Swain & Lapkin, 2001). Furthermore, TBLT supports the development of communicative competence by prioritizing meaning-focused communication while also addressing linguistic and interactional competence (Ellis, 2003). Learners, while concentrating on conveying and comprehending messages, must also attend to linguistic forms for effective learning to occur (Ellis & Shintani, 2014).

TBLT has demonstrated its effectiveness in promoting L2 development, particularly in lexical and grammatical aspects, predominantly in the context of English. For instance, Nuevo et al. (2011)'s study delves into the intricate relationship between task complexity and modified output, examining its connection to L2 grammar development. In this investigation, high complexity tasks prompted more self-repair, which, intriguingly, was linked to improved locative learning. Conversely, low complexity tasks were correlated with more successful past tense acquisition. In another study by Révész and Han (2006), the focus was on understanding the influence of task-related factors on the effectiveness of recasts in second language acquisition. To do so, 36 adult ESL learners were divided into groups, and the impact of task content familiarity and task type on their ability to use the past progressive form was examined. Moving forward, Révész, Sachs, and Hama (2014) undertook an investigation aimed at enhancing the learning of the past counterfactual construction during recasts in communicative activities. The results showed that while input frequency appeared to have no substantial effect, it was the simpler tasks that notably improved oral production gains.

However, when we broaden our perspective to encompass studies beyond grammar and lexis, such as Kim and McDonough's (2008) work, we encounter a different aspect of L2 learning – pronunciation. Here, the focus was primarily within lexis. De Ridder et al.(2007), on the other hand, presented a comprehensive approach, incorporating an independent L2 pronunciation and intonation measure in addition to lexis and grammar. However, it is noteworthy that only one of the four components in their study was task-based, which somewhat limits its contribution to understanding whether task-based designs effectively promote L2 pronunciation accuracy. Even in studies examining oral gains, such as Tonkyn's (2012) work, there was a conspicuous absence of a pronunciation measure. In these cases, L2 accuracy measures, alongside fluency and complexity, were predominantly reserved for lexis and grammar assessment. Consequently, despite the progress made in understanding the efficacy of task-based designs in L2 speech acquisition, the empirical exploration of whether these designs genuinely enhance L2 pronunciation accuracy remains an open question in the field of second language acquisition.

While TBLT theoretically includes pronunciation (Ellis, 2009), and research notes form-focused episodes within meaning-based tasks (Ellis et al., 2011; Gurzynski-Weiss & Baralt, 2014; Loewen, 2005), few studies directly investigate TBLT's impact on pronunciation. Some studies combine TBLT with L2 phonetics/phonology to enhance the understanding of instructed second language acquisition (SLA) and L2 speech development. Loewen and Isbell (2017) explore the influence of modality, learner L1, and task type on pronunciation-related language-related errors (LREs) among L2 English learners, emphasizing these factors' role in pronunciation during task-based interactions. Parlak and Ziegler (2017)) examine how feedback, specifically recasts, affects lexical stress acquisition by Arabic-speaking English learners, highlighting recasts' positive impact. McKinnon(2017) investigates task-based instruction's effects on prosody in Spanish L2 learners, emphasizing its potential in suprasegmental learning. Jung, Kim, and Murphy (2017) apply auditory priming and task repetition to analyze English stress patterns in L2 learners, affirming their efficacy in enhancing prosody. Solon, Long, and Gurzynski-Weiss (2017) explore the impact of cognitively simple or complex dyadic map tasks that emphasize pronunciation's importance, revealing implications of the cognition hypothesis in pronunciation aspects. Gordon's (2021) study examines the combination of explicit pronunciation instruction with TBLT among English-as-a-foreign-language learners, underscoring the need for systematic approaches and various task complexities to improve L2 pronunciation skills.

These studies collectively enhance our understanding of modality, task characteristics, and cognitive factors in shaping pronunciation and language development during task-based interactions, benefiting both theory and language education practice. However, none of these studies has explored the effectiveness of TBLT on segmental perception or examined the potential mediation of personal factors, such as auditory processing abilities, in TBLT effectiveness.

2.3. Auditory Processing

While individuals may receive identical language instruction for the same duration, their L2 proficiency outcomes often diverge significantly (Saito et al., 2017). This variation is attributed to differences in individuals' abilities to notice, elaborate on, and effectively utilize input opportunities – collectively known as 'aptitude.' Unlike mutable factors like motivation, aptitude remains relatively stable across varying L2 learning experiences. Carroll and Sapon's (1959) framework introduced pivotal concepts such as associative memory and phonemic coding ability, culminating in the development of the Modern Language Aptitude Test (MLAT). MLAT scores originally correlated with L2 learners' performance and grades, primarily in form-focused L2 learning contexts. Recent studies, exemplified by Linck et al. (2013), underscore the significance of explicit language learning, working memory, and implicit learning abilities in achieving high L2 proficiency. The LLAMA test, an extension of MLAT, predicts L2 lexicogrammar development, including sound sequence recognition, a critical element in L2 grammatical attainment (Granena, 2013). Furthermore, under L2 learning conditions, explicit learning aptitude (including associative memory, phonemic coding, and grammatical inferencing) emerges as a robust predictor of the effectiveness of explicit instruction and corrective feedback (Yalçın & Spada, 2016; Yilmaz & Granena, 2016). There is a notion that certain abilities may be linked to the acquisition and outcomes of second language (L2) pronunciation. For instance, Reiterer and her research team have demonstrated that phonological working memory plays a significant role in the early stages of acquiring new sounds, as it is closely associated with the activation of specific brain regions such as the left supramarginal gyrus and Broca's area (Reiterer et al., 2011). In essence, aptitude plays a crucial role in comprehending the various achievements of L2 learners.

Until now, extensive research has explored whether, to what extent, and how the acquisition of a second language (L2) can be linked to individuals' language aptitude, as demonstrated by Skehan (2019) and Li (2016). Studies investigating the impact of language aptitude on L2 speech have revealed that L2 learners with higher scores in phonemic coding assessments tend to exhibit superior pronunciation skills at the segmental level, as demonstrated by Saito (2019) and Gordon (2021). However, the complexity of "phonemic coding," the focus of our investigation, encompasses a range of memory and analytical skills and only broadly represents phonological awareness, as highlighted by Skehan (2019). There is limited knowledge regarding the specific components of phonemic coding that contribute to the learning of L2 segments and the perceptual and cognitive abilities that influence the prosodic aspects of L2 pronunciation development, as discussed by Saito et al. (2020). Given the well-established significance of auditory processing in first language acquisition, our current study primarily centers on exploring how variations in auditory processing may impact various aspects of L2 pronunciation development, including segmental, prosodic, and fluency dimensions.

Auditory processing refers to the capacity to accurately represent and remember sound attributes, including nonverbal sounds (Mueller et al., 2012). When learners receive linguistic input, they must encode temporal and frequency patterns, known as auditory acuity, and subsequently retain and incorporate them into their actions, a process referred to as audio-motor integration. These auditory processing abilities are believed to underlie various aspects of language acquisition (Goswami, 2015). In a broader sense, auditory processing can be categorized into two distinct skills: auditory acuity and audio-motor integration. Auditory acuity pertains to one's ability to detect subtle differences in sound characteristics such as frequency and timing. This skill is typically assessed through psychoacoustic tasks where participants discriminate synthesized sounds that vary in a specific acoustic dimension (e.g., formant, pitch, duration). Audio-motor integration, on the other hand, relates to the ability to connect auditory input with motor actions. It is typically assessed through tasks in which participants reproduce sequences of melodies and rhythms (for comprehensive reviews of the auditory precision hypothesis in the context of second language acquisition, see Suzukida & Saito, 2021; Saito & Tierney, 2022).

Scholars have explored auditory processing' relevance in the context of post-pubertal second language (L2) speech acquisition (Mueller et al., 2012). Existing literature indicates a moderate to strong connection between auditory processing and L2 speech acquisition (Shao et al., 2023; Saito & Tierney, 2022). Cross-sectional studies suggest that individuals achieving high L2 speech proficiency typically possess extensive immersion experience and refined auditory processing skills (Kachlicka et al., 2019). Longitudinal research reveals that those with stronger auditory processing tend to make more significant improvements during immersion experiences (e.g., Sun et al., 2021).

In contrast, studies conducted in the context of foreign language classrooms have revealed that individuals with more advanced L2 speaking skills tend to exhibit enhanced audio-motor integration skills, as demonstrated by Saito et al. (2021). However, these advanced speakers may not necessarily possess superior auditory acuity, as indicated by Sun et al. (2021). This suggests that while auditory processing is a crucial factor in L2 speech acquisition, the specific subcomponents of auditory processing, namely acuity and integration, may play distinct roles in facilitating the achievement of advanced L2 speech proficiency.

More empirical studies are needed to explore the mediating roles of auditory processing in diverse learning conditions and participant profiles. (Ruan & Saito, 2023; Shao et al., 2023; Saito, 2023). No empirical studies have ever explored how students with varied AP abilities can differentially benefit from task-based language teaching activities.

Lastly, it's important to acknowledge that the construct validity of the auditory processing measures, such as AXB discrimination, is not entirely clear. This concern has been raised in prior research (Snowling et al., 2018) , indicating that the task format may not solely assess participants' auditory perception skills but may also involve various modality-general cognitive abilities, including working memory and attention control (Saito et al., 2020). To gain a deeper understanding of the specific role of auditory perception in L2 learning, future studies should incorporate not only auditory processing tasks but also cognitive ability assessments. It would be fascinating to investigate whether the link between perception and proficiency remains significant even after statistically accounting for cognitive individual differences. Therefore, the current study has included working memory as another observed factor that might affect the learning outcomes.

2.4. Working Memory and L2 Speech

Working memory, commonly assessed using tasks like the reading span task (Daneman & Carpenter, 1980), is a finite-capacity cognitive system that can handle multiple pieces of information at the same time while doing other tasks. It includes temporary storage, simultaneous processing of incoming data, attentional prioritization, inhibition of irrelevant content, and retrieval of existing knowledge, often through strategic searches of long-term memory (Baddeley, 2000).

Significantly, research findings have highlighted the crucial role of working memory in determining success in achieving proficiency in a second language (Linck et al., 2013). Several studies have indicated that stronger working memory abilities are associated with improved performance in L2 reading comprehension (e.g., Fontanini & Tomitch, 2009; Harrington & Sawyer, 1992). There is substantial evidence supporting the idea that phonological working memory plays a role in language processing to some extent. A substantial body of literature has demonstrated positive correlations between phonological working memory and the acquisition of L2 vocabulary and grammar (e.g., Papagno & Vallar, 1995; Atkins & Baddeley, 1998; Ellis & Sinclair, 1996; French & O'Brien, 2008; Masoura & Gathercole, 2005).

Typically, it is believed that superior cognitive performance can enhance the quality of information intake or assist specific learning mechanisms. For instance, a greater working memory capacity provides learners with more time to process and absorb the input by enabling prolonged access, and improved storage quality can enhance perception accuracy (Goldstone, 1998). The establishment of new and robust phonological representations necessitates learners' engagement with the input. Higher-functioning working memory may aid in the rapid processing of spoken input, allowing for more precise retention of what was heard, and ultimately benefiting phonological development (Darcy et al., 2015). However, previous studies have primarily investigated these factors as potential predictors of overall L2 acquisition rather than focusing on their role in L2 phonological processing.

Until O'Brien et al. conducted their pioneering study in 2007, there was limited attention given to the connection between working memory and the acquisition of a second language in speech. Their study demonstrated a robust correlation between phonological memory capacity and the development of oral fluency in adult L2 learners. Since then, researchers have highlighted the significance of phonological working memory in various aspects of L2 speech development, encompassing complexity (Granena & Yilmaz, 2019), phonological processing, and the perception of L2 sounds (Darcy et al., 2015), as well as pronunciation (Baills et al., 2021).

To date, a significant portion of research in second language acquisition, with a focus on interaction, has delved into the role of working memory (WM) in facilitating awareness during feedback and interaction-driven L2 learning. For instance, one of the earliest studies in this domain, conducted by Mackey et al. (2002), examined how WM capacity moderates task-based interaction learning. Their findings revealed that individuals with higher WM capacity were more proficient at recognizing recasts during conversational interaction compared to those with lower WM capacity.

Subsequently, Mackey et al. (2010) established a favorable association between working memory (WM) and the quantity of adapted output produced during collaborative tasks. Meanwhile, in their study, Mackey and Sachs (2012) observed that older learners with elevated WM capacity displayed enhanced skills in formulating questions during interactive activities. Furthermore, Goo (2012) disclosed that, although recasts and metalinguistic explanations yielded comparable effects on learners' acquisition of the that-trace filter in English, WM played a significant role in mediating the efficacy of recasts. These discoveries strongly imply that executive attention plays a pivotal role in recognizing the impact of recasts (Mackey et al., 2002; Mackey et al., 2010; Mackey & Sachs, 2012; Goo, 2012).

Drawing upon the established notion that working memory significantly influences how individuals respond to corrective feedback in task-based language interactions, as previously discussed by Révész et al. (2011), it becomes imperative to investigate its association with the effectiveness of task-based teaching approaches in improving the way learners perceive and articulate individual speech sounds, known as segments. This exploration aims to shed light on the relationship between working memory capacity and the success of task-based language instruction in enhancing segmental perception skills.

Another rationale for our exploration of working memory within this research stems from the recognition that the utilization of auditory processing assessments not only assesses participants' auditory perception skills but may also draw upon various broader cognitive faculties, such as working memory and attention control, as noted by Snowling et al. (2018). To further investigate the distinct role of auditory perception in second language speech acquisition, forthcoming studies should not only incorporate auditory processing measures but also include cognitive tasks. This broader approach would enable researchers to assess whether the correlation between audio-motor integration and classroom-based second language speech acquisition remains statistically significant even after considering participants' phonological short-term memory and processing speed (Saito et al. 2021).

3. Current Study

3.1. Motivation

The current study is motivated by the relatively limited attention given to L2 speech learning within the framework of TBLT. While some studies have acknowledged learners' consideration of phonological form within larger investigations, there is a dearth of research specifically focused on L2 pronunciation within the TBLT context (Solon et al., 2017).

Moreover, the role of auditory processing in various learning conditions and with diverse participant profiles remains underexplored, highlighting the need for further empirical investigations (Ruan & Saito, 2023; Shao et al., 2023; Saito, 2023b; Saito & Tierney, 2022).

Interestingly, there is a notable gap in empirical research regarding how students with varying auditory processing abilities might differently benefit from task-based language teaching activities. Moreover, this study aims to address these gaps by examining the influence of working memory on L2 speech learning within the context of task-based activities.

Specifically, the goal was to examine the effectiveness of task design on the perception of two difficult vowel contrasts for Chinese EFL learners. The selected phonological contrasts were /e/-/æ/ (e.g., pen vs. pan) and /ɪ/-/iː/ (e.g., ship vs. sheep) which are known to be challenging for Chinese speakers because the absence of the /æ/ sound in Chinese often leads to nasalization and potential confusion with /e/. Additionally, the contrast between /iː/ and /ɪ/ lacks an equivalent in Chinese, further contributing to the complexity of the task (Brown, 1988). Therefore, the current study set out to address the following research questions:

1. To what extent does the task-based instruction can help Chinese learners of English improve their L2 vowel perception (English /e/-/æ/ and /ɪ/-/iː/)?
2. To what extent do individual differences in domain-general auditory processing abilities and working memory mediate the learning outcomes?

4. Method

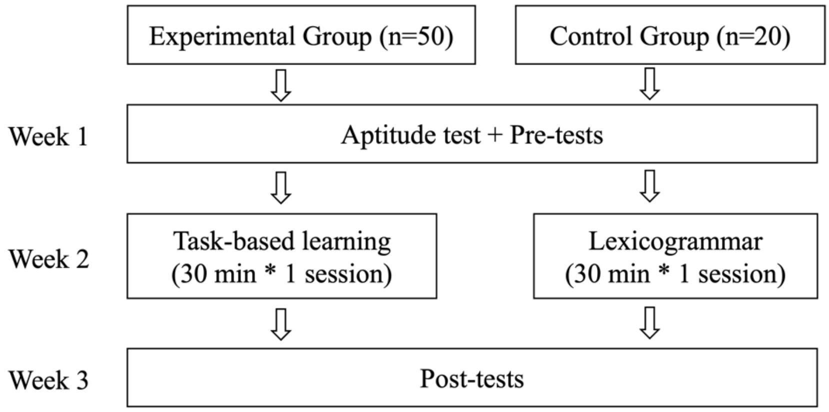
4.1. Design

In this study, a quasi-experimental pre- and post-test design was employed, involving a total of 70 participants who were randomly assigned to either the experimental group (n = 50) or the control group (n = 20). A substantial number of participants were recruited for the experimental group, allowing for subsequent division into subgroups based on high and low auditory processing capabilities. This division aimed to investigate how different aptitude profiles could influence instructional gains. Including the control group was essential to account for potential test-retest effects since identical materials were utilized for both the pre- and post-tests. Despite the logistical constraints, all experiments were conducted via Zoom, with considerable measures taken to ensure training quality. To facilitate interaction and engagement, participants were organized into smaller subgroups of two.

Figure 1 visually summarizes the design of the study. In Week 1, all participants completed pre-tests and aptitude assessments. In Week 2, the experimental group engaged in a 30-minute task-based instructional session, while the control group received lexicogrammar training. Subsequently, in Week 3, all participants completed the post-tests.

During both the pre- and post-test phases, participants completed a forced-choice identification task. This task aimed to assess the influence of instruction on the participants' perception abilities concerning English vowel pairs, specifically /e/ - /æ/ and /ɪ/ - /i:/, in both trained and untrained lexical contexts. Participants' perception abilities were assessed via a forced-choice identification task.

Figure 1. *Summary of Research Design.*



4.2. Participants

The study involved young adult Chinese English-as-a-Foreign-Language (EFL) learners hailing from Fujian province in southeastern China. These learners were native Mandarin Chinese speakers. To recruit participants, advertisements were circulated among students via online channels and email. Interested individuals reached out to the researcher, expressed their willingness to participate, and completed the necessary consent forms, also scheduling their pre-test appointments. The study comprised a total of 70 young Chinese EFL learners, consisting of 25 males and 45 females. The average age across the entire sample was 18 years, with an age range spanning from 17 to 20. All participants had previously undertaken a vocabulary test on the VocabularySize.com website, revealing an average vocabulary size of 6158 word families, ranging from 3800 to 10500. Despite having substantial prior EFL learning experience, such as over 5 years, none had experienced full immersion in an English-speaking environment. All participants confirmed having normal hearing abilities. Given their location in China, opportunities for interaction with native and non-native English speakers were notably scarce. Following the pre-tests, participants were randomly allocated to either the experimental group (n = 50) or the control group (n = 20). As outlined in the Results section, the experimental group was further subdivided into two subgroups based on their auditory processing and working memory capacities. Notably, none of the participants reported any hearing impairments.

4.3. Instructional Treatment

Instructor.

The teacher for both sets of students was a highly proficient English speaker who was a native Mandarin speaker. This instructor possessed considerable experience in teaching English in China for over four years and held a Master's degree in Teaching English to Speakers of Other Languages (TESOL)..

Target of instruction.

The study focuses on two phonological contrasts, English /e/ and /æ/ (e.g., pen vs. pan) and English /ɪ/ and /i:/ (e.g., ship vs. sheep). According to the “Ranking Ordering of RP Phoneme Pairs Commonly Conflated by

Learners", /e/-/æ/ ranked 10 and /ɪ/-/i:/ ranked 8 in a 10-points scales of importance which means these two pairs of phonemes is of great importance for teaching (Brown, 1988).

English vowels /ɪ/ and /i:/ differ in articulation, duration, and spectral characteristics. Native speakers often rely on formant values (Hz) as the primary cue, particularly in General American English (Gottfried, Miller, Payton, 1990). However, in cases of unclear cues, they also consider phonemic length (ms) (Hillenbrand, Clark, & Houde, 2000). In Received Pronunciation, there is a distinct long-short vowel length contrast (/ɪ/ and /i:/) (Ladefoged & Johnson, 2014). L1 Mandarin learners of English often struggle to differentiate between /ɪ/ and /i:/, tending to categorize both as /i:/ (Wang & Munro, 2004). They may overemphasize duration over spectral cues, particularly in General American.

In line with the perceptual assimilation model (Best, 1995; Best & Tyler, 2007), when considering that Mandarin's vowel inventory only includes English /i:/, L1 Mandarin learners of English initially tend to perceive English /ɪ/ and /i:/ as equivalents to the Mandarin /i:/ English /æ/ may be perceived as highly similar to the Chinese /e/ sound due to their shared front and unrounded vowel characteristics, primarily differing in tongue height. This similarity can hinder the establishment of a distinct phonetic category for English /æ/ (Speech Learning Model, Flege, 1995), resulting in pronunciation difficulties and a pattern of substitution, as observed in Chan's study (2010). In Chan's research, it was found that participants exhibited significant inaccuracies in producing /æ/, with an accuracy rate of only 16.4%. Notably, 97.7% of participants mispronounced /æ/ as /e/, while 30.7% of them mispronounced /e/ as /æ/. This challenge stems from the absence of the /æ/ sound in Chinese, often leading to nasalization and potential confusion with /e/ (Swan & Smith, 2001).

The mispronunciation of /e/-/æ/ and /ɪ/-/i:/ can impede successful communication due to the presence of minimal pairs and the functional load principle. Therefore, accurate pronunciation of these two pairs of vowels is crucial (Brown, 1988; Best & Tyler, 2007;; Munro & Derwing, 2006; Suzukida & Saito, 2021).

Treatment.

Regarding the experimental group, participants underwent a 30-minute task-oriented training session, focusing on negotiating a summer holiday travel plan. The instructional materials were initially designed based on the work by Mora (2021), which included various context-driven tasks aimed at assisting native Catalan/Spanish speakers in learning English phonemes /iː-ɪ/ and /æ-ʌ/. These materials were subsequently modified to cater to Chinese learners' acquisition of English /e/-/æ/ and /ɪ/-/i:/ phonemes.

A set of 20 minimal pairs, which only differed in terms of /e/ - /æ/ and /ɪ/ - /i:/ sounds (refer to "trained items" in Table 1), were incorporated into the activity. Initially, participants were involved in an awareness-raising exercise lasting 5 minutes. During this phase, they heard one of the words from the minimal pairs twice and indicated the corresponding image. Subsequently, they listened to sentences containing the target words and verified their responses. Following this, participants participated in a comprehension task involving listening to a conversation. They were informed that they would be expected to engage in dialogues akin to this conversation in the primary task. Afterward, they proceeded to the primary task, which spanned 25 minutes. The primary objective of this task was to assist students in enhancing their accuracy in pronouncing English /e/ - /æ/ and /ɪ/ - /i:/ sounds through the process of exchanging information and making decisions.

Additionally, the tasks were designed with the intention of affording opportunities for communication using specific linguistic features, as proposed by Ellis (2009). In particular, learners needed to be able to differentiate between L2 vowel distinctions (/e/ - /æ/ and /ɪ/ - /i:/) in order to successfully complete the task (i.e., task indispensability; Loschky & Bley-Vroman, 1993).

In order to enhance the salience of the specific words, they were presented in bold text, which is a typographical emphasis. Throughout the activity, participants were instructed to remain primarily focused on the meaning, specifically the discussion regarding the travel plan, while also paying attention to distinctions between English phonemes /ɪ/ - /iː/ and /e/ - /æ/. In this task, students worked in pairs, and the teacher observed their performance. Each participant had unique preferences and personal circumstances that their partners were unaware of. The task required students to engage in a conversation with their partners, share information, and collectively choose three destination preferences based on individual preferences and conditions. Subsequently, they were expected to present their travel plan. The teacher made sure students were using the L2 and promoted spontaneous talk and confidence building. For the details of the task that were used in class, see Supporting Information B.

As for the control group, participants spent a similar amount of time (30 minutes * 1 session) practicing English vocabulary and grammar with the researcher. The materials largely consisted of supplementary exercises related to the content covered in their regular school EFL lessons. These sessions did not involve any oral communication practice or specific training. The inclusion of the control group aimed to discern any potential test-retest effects, as identical materials were employed for both the pre-test and post-test assessments. Since all the study participants resided in China and had access solely to grammar-focused classroom English instruction, none of them indicated receiving any additional pronunciation training involving English /e/-/æ/ and /ɪ/-/iː/ between the pre-test and post-test stages.

Table 1. *Forty Tokens in the Perception Tests*

Trained Items	Untrained Items
1. and – end	1. bag-beg
2. band – bend	2. beat – bit
3. bean – bin	3. dad-dead
4. cheap – chip	4. feel – fill
5. gas – guess	5. had-head
6. marry – merry	6. land – lend
7. meal – mill	7. man-men
8. pan – pen	8. seat – sit
9. sheep – ship	9. sick – seek
10. steel – still	10. slip – sleep

* *Note*. The lexical items are listed alphabetically in the table. They were presented in random order in the tests.

* They all fall among the first 2000 most common word families according to BNC/COCA Word Frequency List (Nation, 2012)

4.4. Measure of L2 Speech Development

In order to assess the influence of task-based instruction on how participants perceived and acted upon specific phonetic structures in English (/iː – ɪ/ and /æ – e/), a forced-choice identification task was employed for both the initial and final assessments.

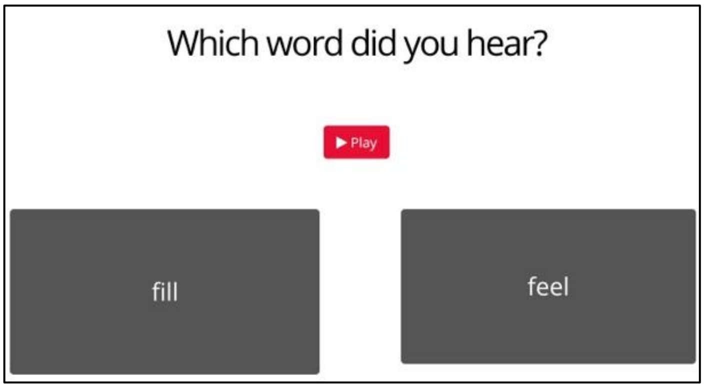
Materials.

The testing materials comprised both trained and untrained items. The untrained words were employed to assess the degree to which instructional effects could be applied to unfamiliar words. Within the testing materials (refer to Table 1), there were a total of 40 target tokens, all belonging to the first 2,000 most common word families as per the BNC/COCA Word Frequency List (Nation, 2012). To minimize the influence of word frequency and familiarity on test performance, it's noteworthy that the learners in this study possessed word families ranging from 3,800 to 10,500. All speech samples were generated using online tools and adhered to the standards of British English.

Procedure.

After receiving instructions from the researcher, participants used their personal computers equipped with headphones to access the test materials available on the online platform Gorilla (Anwyl-Irvine, Massonnie´, Flitton, Kirkham, & Evershed, 2020). A total of 80 stimuli, consisting of 40 trained and untrained target words spoken by two native speakers, were presented to them in a randomized sequence. For each stimulus, participants were required to determine the word they heard by selecting one of two written options displayed on the computer screen (refer to Fig. 2). Their performance was assessed and recorded on a 100-point scale for accuracy.

Figure 2. *A Screenshot of the Perception Test*



4.5. Measures of Auditory Processing

Auditory acuity (formant, pitch, duration).

In line with prior research (e.g., Kachlicka et al., 2019), this study employed three distinct AXB discrimination tasks to evaluate participants' auditory sensitivities across various dimensions, encompassing fundamental frequency (pitch), formants, and duration perception. These tasks utilized synthesized, nonverbal stimuli with simplified acoustic characteristics, rendering them unperceivable as speech by individuals with normal hearing (for listener judgments of these nonverbal stimuli, refer to Saito, Kachlicka, et al., 2022; for a comprehensive description of the auditory processing tests, consult Saito & Tierney, 2023).

In each trial, participants were presented with three non-verbal sounds and tasked with identifying the one (either the first or third) that differed from the other two by selecting the corresponding number ("1" or "3") using a mouse (see Fig. 3). Given that each sound differed in just one acoustic parameter (pitch, formants, or duration), each task aimed to measure the smallest discernible difference participants could perceive in the target dimension.

For each test, custom MATLAB scripts were utilized to generate one reference stimulus and 100 target stimuli. In the pitch discrimination assessments, the stimuli consisted of 250-ms-long complex tones with a fundamental frequency (F0) of 330 Hz. The reference stimulus maintained an F0 of 330 Hz, while the target stimuli ranged from 330.3 to 360 Hz in 0.3-Hz increments. In the case of duration discrimination, the reference stimulus had a length of 250 ms, and the target stimuli ranged from 252.5 ms to 500 ms in 2.5 ms steps.

Complex tones with an F0 of 100 Hz and comprising three formants (F1 = 500 Hz, F2 = 1500–1700, F3 = 2500 Hz) were employed for formant discrimination. The reference stimulus had an F2 set at 1500 Hz, whereas the F2 of the target stimuli varied from 1502 Hz to 1700 Hz in 2 Hz increments. Utilizing Levitt’s (1971) adaptive up-down procedure, the test initially commenced at Level 50 (out of 100 levels) and automatically adjusted the difficulty level based on the participant’s performance. Difficulty increased by 10 steps after three consecutive correct responses and decreased by 10 steps after a single incorrect response. The step size reduced to five after the first reversal, then to two after the second, and finally to one after the third, continuing until task completion. Participants' ultimate scores were recorded on a 100-point scale, with lower scores indicative of more precise auditory processing concerning formants, pitch, and duration.

All training and assessment materials were accessed on the open science platform for potential replication in the future, L2 Speech Tools (http://sla-speech-tools.com/).

Figure 3. *A Screenshot of Auditory Processing Discrimination Tasks*

Audio-motor integration (rhythm, melody)

Drawing upon the protocols and materials developed by Tierney et al. (2017), Saito, Suzuki, et al. (2021), and Sun et al. (2021), two distinct tasks involving audio-motor integration were administered, specifically focusing on rhythm and melody reproduction.

For the rhythm reproduction task, ten distinct rhythmic patterns, each lasting 3.2 seconds per sample, were created and presented to the participants. These rhythms were selected from those utilized in the rhythmic patterns designed by Povel and Essens (1985). Each rhythmic pattern consisted of a sequence of 16 segments, each lasting 200 milliseconds. Segments containing a drum hit featured a conga drum sound sourced from freesound.org within the initial 150 milliseconds, while segments designated as "rest" contained no sound. In each trial, participants listened to the same stimulus three times and were subsequently required to replicate the rhythmic sequences by pressing the space bar (see Fig. 4). The time intervals between their space bar presses were adjusted to the nearest set interval (200, 400, 600, 800 milliseconds). The accuracy of their responses was then determined by calculating the ratio of correctly reproduced hits and rests in 200-millisecond intervals compared to the target stimuli.

For the melody reproduction task, ten melodies were created, each comprising a sequence of seven notes, with each note lasting 300 milliseconds. These notes were drawn from a five-note harmonic complex tone scale with fundamental frequencies of 220, 246.9, 277.2, 311.1, and 329.6Hz, corresponding to the first five notes of the A major scale. The melodies were generated by initiating them on the third tone (277.2Hz) of the scale and subsequently selecting the next note to be either one note higher (246.9Hz) or lower (311.1Hz) than the previous one, creating a sequence of seven notes. If a note reached the lower (220Hz) or upper (329.6Hz) limits of the scale, the subsequent note was chosen to be closer to the center of the scale or identical to the previous note. Participants listened to these melodies three times each and were tasked with reproducing them by selecting five buttons labeled from "5" to "1" in descending order (see Fig. 5), corresponding to the highest to the lowest tones. The accuracy of their reproductions was assessed by comparing their initial seven button presses to the original melody, and the mean accuracy ratio was computed across all ten melodies.

Figure 4. *A Screenshot of Rhythm Reproduction Task.*

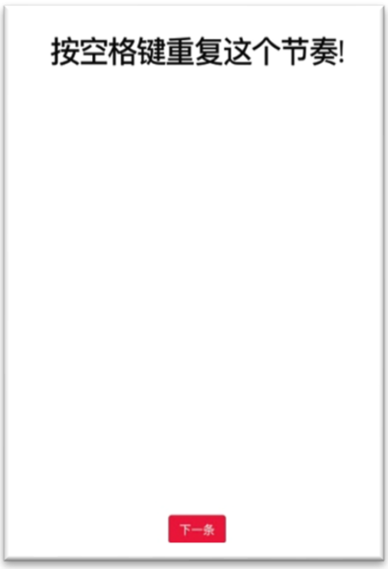
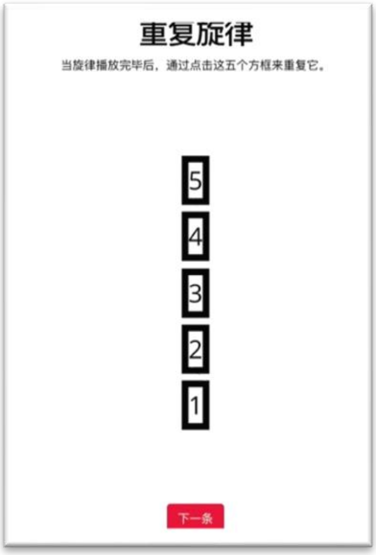


Figure 5. *A Screenshot of Melody Reproduction Task.*

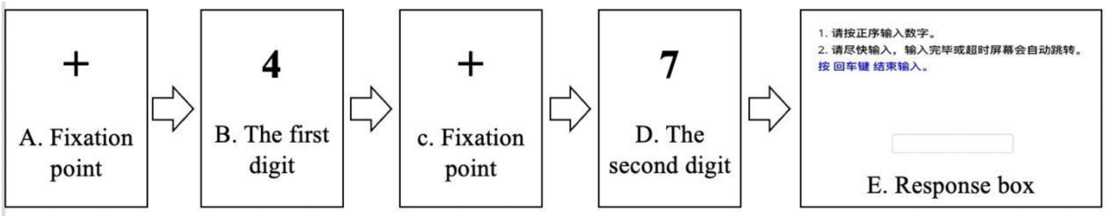


4.6. Measures of Working Memory

Working memory was assessed using the visual and text-entry digital span tests, which were adapted and customized based on Dean's work (2020). It was originally designed for diagnosing children's learning difficulties related to verbal working memory. During these tests, pre-recorded sequences of numbers were displayed on the screen at a rate of one digit per second, with a brief fixation point in between digits (lasting 100ms). Participants were instructed to recall the numbers by entering them into an input box once each sequence concluded (see Fig. 6). Responses were obligatory, and participants needed to press the "return" key on the keyboard to confirm their response and proceed to the next trial.

The digit span forward test (DSF) was employed to evaluate participants' phonological short-term memory (PSTM) capacity. In this task, participants were required to recall the numbers in the exact order they were presented. Conversely, the digit span backward test (DSB) assessed complex working memory processing, where participants had to input numbers in the reverse order they were displayed. The DSF included nine spans, ranging from 2 digits to 10 digits, while the DSB consisted of eight spans (from 2 digits to 9 digits). Each span comprised two trials, with successful completion of at least one trial leading to the next level (one digit longer than the previous one) until participants were unable to correctly recall either of the two trials of the same length. The working memory capacity was determined by the longest spans participants successfully entered without error. The two tests were administered sequentially, with the DSF preceding the DSB. An optional break between the tasks was provided, and this segment typically took approximately 5 to 10 minutes to complete (Dean, 2020).

Figure 6. *A Screenshot of a Two-digit Span Test*



5. Results

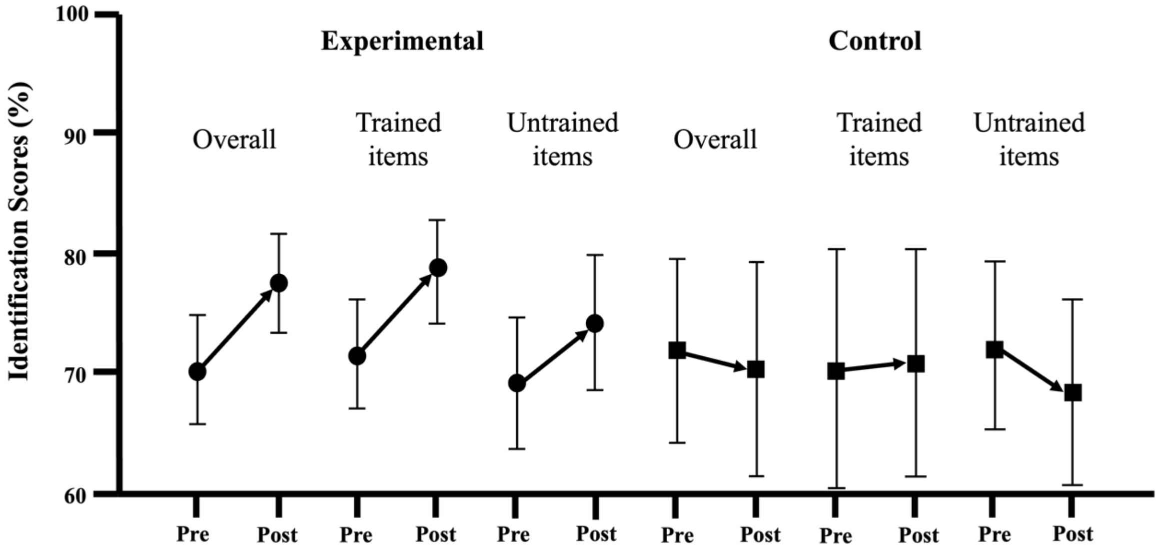
5.1. Overall Improvement (Pre-to Post-Tests)

Table 2 and Figure 7 present pre-test and post-test correct identification scores (%) under three conditions: overall (n = 40 items), trained (n = 20 items), and untrained (n = 20 items). The normality of pre-test scores for both experimental and control groups was assessed using a Kolmogorov-Smirnov test, which revealed no significant deviation from normal distribution ($p > .05$). To investigate potential pre-existing differences in perceptual accuracy of target sounds (English /iː – ɪ/ and /æ – e/), independent t-tests were conducted on total pre-test scores. The results showed no significant between-group differences at the pre-test stage ($t = -.129$, $p = .898$, $d = -.034$), indicating comparable vowel perception performance before the treatment.

Table 2. *Summary of L2 Vowel Identification Scores*

	Experiment (n = 50)				Control (n = 20)			
	M	SD	95% CI		M	SD	95% CI	
			Lower	Upper			Lower	Upper
L2 vowel identification (%)								
Overall (pre)	70.7	16	66.1	75.3	71.2	16	64.0	78.5
Trained (pre)	72.0	16	67.4	76.7	70.5	21	60.1	80.3
Untrained (pre)	69.0	21	63.1	74.9	72.0	15	65.1	78.9
Overall (post)	77.1	14	73.0	81.1	70.3	17	62.2	78.3
Trained (post)	78.2	14	74.1	82.3	71.5	20	62.2	80.8
Untrained (post)	74.4	19	68.9	79.9	68.5	18	60.3	76.7

Figure 7. *95% Confidence Intervals and Mean Values of the Learners' Perception Scores.*



To investigate the impact of task-based instruction on segmental perception development, we conducted a three-way ANOVA with Group (Experimental vs Control) as a between-subjects factor, and Lexis (trained vs. untrained) and Time (pre-/post-tests) as within-subjects factors. We found a significant Group × Time interaction effect, $F(1, 68) = 6.603$, $p = .012$, $\eta p^2 = .089$.

The multiple comparisons revealed that the control group's performance showed no statistically significant change over time (M = 71.2% → 70.3%, $p = .592$, $\eta p^2 = 0.004$), indicating the absence of a test-retest effect in our study.

Conversely, the experimental group demonstrated a significant overall improvement in scores from pre- to post-tests (M = 70.7% → 77.1%, $p < .001$, $\eta p^2 = 0.187$). In line with Cohen's (1988) benchmarks, this effect size can be considered medium to large.

However, the three-way Group × Time × Lexis interaction did not reach statistical significance, $F(1, 68) = 0.377$, $p = .541$, $\eta p^2 = 0.006$.
Notably, we observed a significant main effect for Lexis, indicating that learners in the experimental group performed significantly better on both trained (M = 72% → 78.2%, $p = .014$, $\eta p^2 = 0.085$) and untrained (M = 69% → 74.4%, $p = .005$, $\eta p^2 = 0.113$) items compared to the control group.

Collectively, the outcomes suggest that the experimental group exhibited notable enhancements in their capacity to differentiate between English /iː – ɪ/ and /æ – e/ sounds, irrespective of the word context, including both trained and untrained instances. These findings underscore their ability to apply their acquired skills broadly. The gains here were not due to test-retest effects.

5.2. Language Aptitude and L2 Speech Learning

The subsequent step in the statistical analysis aimed to investigate whether the enhancement observed in the experimental group could be attributed to the participants' auditory processing abilities and working memory.

Following the procedures outlined in prior literature (e.g., Kachlicka et al., 2019), a composite auditory acuity score was computed for the Experimental group by averaging their formant, pitch, and duration discrimination scores. Similarly, a composite integration score was calculated by averaging their rhythm and melody reproduction scores.

The normality of the data was assessed using Kolmogorov-Smirnov tests. Auditory acuity scores and working memory scores exhibited distributions that were comparable to a normal distribution ($p > .05$). However, audio-motor integration scores showed a departure from normality, as indicated by a

Kolmogorov-Smirnov test (p = .006). Nevertheless, it is noteworthy that, when evaluating normality using skewness (-0.125) and kurtosis (-1.351) in conjunction with Q-Q plots for audio-motor integration scores, no significant skewness was observed in the data.

Consistent with previous research on the interaction between aptitude and treatment effects, we conducted two distinct analyses."

Variance-Based Analyses.

Given the considerable variability in auditory processing abilities among learners, we examined whether participants' gains were associated with their individual auditory profiles while controlling for their pre-test performance. To accomplish this, we conducted partial correlation analyses. This approach was chosen to account for the potential influence of participants' pre-test scores on the relationship between raw gains and auditory scores (e.g., learners with lower pre-test scores may exhibit larger gains due to greater room for improvement).

The correlation coefficients revealed no statistically significant relationships between gains and auditory acuity scores (r = -0.074, p = .613), audio-motor integration scores (r = 0.098, p = .502), and working memory scores (r = 0.096, p = .511).

For a more detailed investigation, we conducted a partial correlation analysis, focusing on specific subsets of auditory acuity processing (pitch, formant, duration) and audio-motor integration ability (melody and rhythm reproduction). This analysis revealed a significant negative relationship between participants' gains and their formant scores (r = -0.287, p = .045). Notably, formant scores indicate the minimum difference that participants could hear for formant frequency. Therefore, lower scores reflect better discrimination ability among participants. The negative relationship found suggests that those with greater discrimination ability in formant processing exhibited larger gains.

Mean-Based Analyses.

Certain researchers have contended that, although adult second language (L2) learners have the capacity to acquire novel sounds, this progress may be constrained to specific individuals facing auditory challenges (Perrachione et al., 2011). In our investigation, we employed a group-level analysis to explore whether individuals with contrasting levels of aptitude experienced distinct advantages from task-based instruction. We implemented a median-split methodology akin to the approach adopted by Chandrasekaran et al. (2010). When the experimental participants were categorized based on their formant discrimination scores, they were split into two distinct subgroups within the experimental group, resulting in a high-formant subgroup (consisting of 24 individuals; mean = 19.20, standard deviation = 8.95, range = 5.80 to 33.83) and a low-formant subgroup (comprising 26 individuals; mean = 49.28, standard deviation = 9.02, range = 36.00 to 70.83) as detailed in Table 3. The participants' scores for vowel perception were then subjected to a two-way repeated-measures ANOVA, with Group (high-formant group, low-formant group, and control group) as a between-group factor and Time (pre vs. post) as a within-group factor. The analysis yielded significant main effects for Time, indicating a difference, $F (1, 67) = 9.917$, $p = .002$, $\eta p^2 = .129$. Furthermore, the interaction effects between Group and Time were statistically significant, $F (2, 67) = 3.924$, $p = .024$, $\eta p^2 = .105$.

After undergoing the intervention, participants in the high-formant group exhibited substantial improvements in their accuracy scores over time, demonstrating significant and sizeable effects (from an average of 73.9% to 81.7%, $F (1, 67) = 6.154$, $p < .001$, $\eta p^2 = .162$). Furthermore, in post-tests, they surpassed participants with lower auditory scores in a statistically meaningful manner (81.7% compared to 72.9%, $F (1, 67) = 13.145$, $p = .014$, $\eta p^2 = .087$). Notably, the control group's perception scores remained unchanged between the pre- and post-tests ($p = .672$). For a visual representation of group performance, please refer to Figure 8.

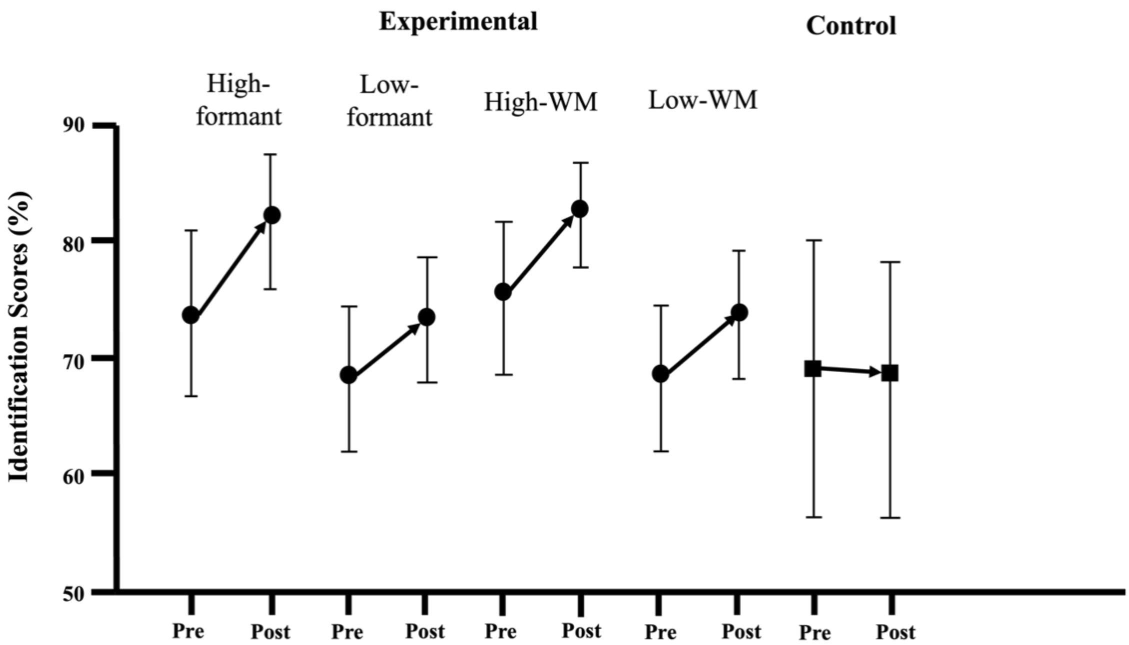
When we categorized the experimental group based on their working memory scores, we divided the total of 50 participants within the experimental group into two subgroups: one with high working memory (WM) scores ($n = 28$; mean = 20.36, standard deviation = 3.346, range = 17 to 28), and another with low working memory (WM) scores ($n = 22$; mean = 12.79, standard deviation = 2.455, range = 6 to 16) as detailed in Table 3. The participants' scores in vowel perception underwent a two-way repeated-measures ANOVA, considering Group (comprising the high-WM group, low-WM score group, and control group) as a factor between groups and Time (pre vs. post) as a factor within groups. This analysis revealed significant main effects of Time, $F(1, 67) = 9.817$, $p = .003$, $\eta p^2 = .128$. Moreover, the Group × Time interaction effects reached statistical significance, $F(2, 67) = 4.213$, $p = .032$, $\eta p^2 = .097$.

Following the treatment, participants with high working memory (WM) scores exhibited significant and substantial improvements in their accuracy scores over time (Mean = 74.8% to 81.8%), as indicated by a statistically significant effect ($F(1, 67) = 6.154$, $p = .003$, $\eta p^2 = .128$). Moreover, they outperformed participants with low WM scores in the post-tests (Mean = 81.8% vs. 73.4%), a difference that reached statistical significance ($F(1, 67) = 13.145$, $p = .004$, $\eta p^2 = .116$). In contrast, the control group did not demonstrate any significant improvement in perception scores from the pre-test to the post-test ($p = .672$). For a visual representation of group performance, please consult Figure 8.

Table 3. *Analysis of L2 Vowel Recognition Performance in Relation to Formant and Working Memory Scores Compared to a Control Group.*

	Aptitude scores			Pre-test (%)				Post-test (%)				Gains (%)			
						95% CI				95% CI				95% CI	
	M	SD	Range	M	SD	Lower	Upper	M	SD	Lower	Upper	M	SD	Lower	Upper
High-formant (n = 24)	19.20	8.95	5.80-33.83	73.9	16	67.1	80.9	81.7	14	75.8	87.5	7.7	11	3.0	12.4
Low-formant (n = 26)	49.28	9.02	36.00-70.83	67.7	16	61.2	74.2	72.9	14	67.4	78.4	5.1	11	.7	9.7
High-WM (n = 28)	20.36	3.35	18.88-21.85	74.8	15	68.3	81.2	81.8	11	76.9	86.7	7.0	12	1.5	12.5
Low-WM (n = 22)	12.79	2.46	6.00-16.00	67.5	17	60.9	74.1	73.4	16	67.4	79.4	5.8	10	1.9	9.8
Control (n = 20)				68.5	17	56.5	80.5	67.5	16	56.4	78.6	-1	9	-7.7	5.7

Figure 8. *Analysis of Learners' Vowel Identification Scores: Mean Values and 95% Confidence Intervals in Relation to Formant and Working Memory Conditions Compared to a Control Group*



6. Discussion

6.1. TBLT and L2 Speech Teaching and Learning

The first research question asked how task-based instruction help Chinese learners of English improve their L2 vowel acquisition (English /e/-/æ/ and /ɪ/-/iː/). The findings of this study offer several noteworthy conclusions. Initially, there were no significant differences between the experimental and control groups in their L2 vowel perception abilities, as indicated by comparable pre-test scores, highlighting the equivalence of the two groups at the outset. The task-based instructional approach, lasting only 30 minutes, had a remarkably positive impact on L2 vowel perception among learners in the experimental group. These learners exhibited significant improvements of 9.05% overall, comprising specific gains of 8.6% for trained words and 7.8% for untrained words. Furthermore, these improvements were associated with medium-to-large effect sizes ($\eta p^2 = 0.187$), emphasizing the substantial pedagogical benefits of the task-based instruction. Importantly, the stability of the control group's performance over time ($M = 71.2\% \rightarrow 70.3\%$, $p = .592$, $\eta p^2 = 0.004$) suggests that the improvements observed in the experimental group were not merely a result of test-retest effects. In summary, the study's results indicate that a brief 30-minute task-based teaching session had a substantial and statistically significant impact on improving L2 vowel perception, with effects comparable to or exceeding those reported in longer-duration training sessions in previous research.

These findings carry significant implications for language pedagogy, particularly regarding the effective enhancement of L2 segmental development through carefully designed tasks that prioritize the target segments' essential role. However, it is important to emphasize that sustainable, long-term progress necessitates the consistent integration of such instructional strategies in language classrooms over an extended duration. These results echo previous literature supporting the substantial impact of task-based instruction on promoting the acquisition of L2 speech at the segmental level, as demonstrated in prior research.

The results also revealed that participants who received task-based instruction showed improvement not only in recognizing or performing tasks related to the trained items (those explicitly taught during the instruction) but also in tasks related to untrained items (those that were not explicitly taught during the instruction). This finding suggests that the skills or knowledge gained through the instructional intervention had a broader application beyond the specific items that were part of the training. This demonstrates the potential effectiveness of the instructional approach in promoting generalization and transfer of skills, indicating that learners are not merely memorizing or rote learning specific content but are acquiring more robust and flexible skills that can be applied to a wider range of contexts.

6.2. Auditory Processing and Working Memory and L2 Speech Learning

The findings derived from partial correlation analyses in this study reveal that there were no statistically significant associations between the participants' overall auditory acuity scores, audio-motor integration scores, or working memory scores and the extent of their gains in L2 speech perception. These results suggest that, within the specific context of this investigation, the participants' improvements in L2 speech perception were not strongly contingent upon their general auditory processing abilities or working memory.

However, a more nuanced exploration focused on distinct facets of auditory acuity processing, particularly formant discrimination, unveiled a noteworthy pattern. Specifically, a significant negative correlation emerged between participants' gains in L2 speech perception and their formant scores ($r = -0.287$, $p = .045$). Formant scores in this context serve as proxies for participants' proficiency in distinguishing subtle differences in formant frequencies, with lower scores indicative of heightened discrimination capabilities.

This negative correlation signifies that individuals possessing superior formant discrimination abilities, denoting a heightened capacity for distinguishing formant frequency disparities, exhibited more substantial gains in L2 speech perception following the instructional intervention. Further dissection of the experimental group based on their formant discrimination scores delineated two distinct subgroups: the high-formant group and the low-formant group. Notably, the high-formant group displayed remarkable enhancements in their accuracy scores over time, evincing substantial effects ($p < .001$, $\eta p^2 = .162$).

These findings underscore the pivotal role played by formant discrimination abilities in mediating the efficacy of the instructional intervention. Learners endowed with superior formant discrimination skills not only manifested heightened performance improvements but also surpassed their counterparts with lesser formant discrimination aptitude in the post-intervention assessments. Learners endowed with superior formant discrimination abilities exhibit heightened responsiveness to the intervention which can be attributed to their enhanced aptitude for tracking and retaining acoustic information, enabling them to focus on primary acoustic features crucial for segmental accuracy, particularly high-frequency spectral information. The instructional task in the current study is intentionally crafted to engage participants' ability for tracking and deciphering diverse acoustic dimensions within speech signals, necessitating their successful task completion. This task explicitly mandates participants to differentiate between sounds characterized by variations in spectral attributes, such as formant shapes. In such contexts, learners not only discern novel patterns in spectral and temporal information, often not their primary cues for distinguishing and identifying L1 phonological distinctions (as exemplified by Japanese speakers in English /r/-/l/ acquisition, as observed in Saito & Brajot, 2013), but also exhibit adaptability by reconfiguring their perceptual strategies. This adaptative process entails shifting reliance from certain cues, such as pitch, to others, such as durational information, as evidenced in the case of Chinese speakers acquiring English prosody (Jasmin, Sun, & Tierney, 2020).

Consequently, learners possessing specific auditory skills related to formant frequency discrimination can effectively accommodate and recalibrate their perceptual strategies when confronted with novel acoustic patterns during the instructional task. This adaptability augments their responsiveness to the task, elucidating the criticality of these auditory skills in determining the instructional intervention's effectiveness. The ability to detect and process formant frequencies significantly contributes to learners' capacity to discern and adapt to these novel patterns, ultimately culminating in enhanced segmental accuracy in speech perception and production.

The finding is in line with previous research, as significant interaction effects of time and auditory processing (specifically, formant discrimination) were observed concerning segmental and word stress aspects. These results underscore the pivotal role of auditory processing in influencing the extent of improvement, particularly in the realm of accuracy within L2 pronunciation proficiency (Saito et al., 2020). In the context of L2 speech acquisition, Lengeris and Hazan (2010) conducted a study with Greek speakers, revealing that those with superior formant discrimination abilities not only exhibited higher initial proficiency in L2 English segmentals but also experienced more substantial gains when subjected to auditory training focused on L2 English vowels. This finding is consistent with the work of Chandrasekaran, Kraus, and Wong (2012), who employed neural methods to explore similar phenomena. Moreover, the result aligns with Ruan and Saito (2023)'s study, which suggests that the aptitude-instruction link was dichotomous rather than linear. Specifically, significant improvement resulting from task-based instruction was evident among those with high formant discrimination ability but not among those with low formant discrimination ability. The latter group may have struggled to encode and master the acoustic properties of the target sounds, particularly when their primary focus was on conveying meaning. Low-aptitude learners may not have effectively noticed, remembered, and internalized the L2 input for long-term representations.

Extending the existing body of research on aptitude-treatment interactions (as illustrated, for instance, in the work by Chandrasekaran et al., 2010, concerning the role of auditory processing in explicit L2 speech training), our current investigation introduces an additional dimension by highlighting the role of domain-general auditory processing in mediating the efficacy of meaning-oriented, communicatively authentic instructional approaches. Previous studies have reported noteworthy aptitude effects in the realm of domain-general auditory processing, primarily within naturalistic L2 learning contexts, as opposed to classroom settings (e.g., as observed in the research conducted by Saito et al., 2021). Our study extends this evidence base by establishing a connection between formant discrimination ability and learning outcomes in instructed language learning. These findings provide valuable insights into the intricate relationship between distinct aptitude profiles and their varying degrees of relevance across different phases and types of L2 learning, aligning with the argument advanced by Skehan (2016).

However, two other dimensions of auditory processing appear to have no significant impact on the final level of language attainment. One plausible interpretation of these findings is rooted in the participants' native language background, specifically Mandarin Chinese. Previous studies have indicated that speakers of tonal languages tend to exhibit more robust brainstem representations of fundamental frequencies in comparison to non-tonal language speakers (Bidelman et al., 2011) and demonstrate greater accuracy in pitch discrimination (Giuliano et al., 2011). Consequently, Mandarin L1 speakers may possess a generally heightened capacity for precise pitch perception, enabling them to extract prosodic nuances from speech more effectively than speakers of other languages. This phenomenon may lead to weaker associations between pitch perception and prosodic or phonetic processing. To further explore this hypothesis, future research could investigate the correlations between auditory processing and pronunciation in individuals from diverse linguistic backgrounds.

Similarly, when dividing the experimental group based on WM scores, two subgroups emerged: the high WM scores group and the low WM scores group. The high WM scores group showed a substantial improvement in accuracy scores after the instructional intervention, with large effects (p

= .003, ηp^2 = .128). They also outperformed participants with lower WM scores in the post-tests.

The Cognition Hypothesis posits that language learning tasks with greater cognitive demands can lead to more effective learning outcomes by promoting increased attention, memory engagement, and improved differentiation of language use (Robinson, 2003). This theoretical framework suggests that learners who possess higher cognitive capacity, as indicated by their working memory scores, may be better equipped to handle the cognitive demands of such tasks. The results of current study reveals that learners with higher working memory scores tend to benefit more from task-based instruction. In practical terms, this means that individuals with greater cognitive resources are more adept at navigating the complexities of task-based language learning, resulting in enhanced language proficiency and more successful learning experiences. This alignment between theory and practice underscores the importance of considering learners' cognitive abilities when designing instructional approaches, emphasizing that tailored task-based instruction can be particularly advantageous for those with higher working memory capacities in the pursuit of language acquisition.

Up until now, the majority of research in the field of interaction-oriented Second Language Acquisition (SLA) has primarily focused on understanding the significance of Working Memory (WM) in processes related to the perception of feedback and interaction-driven L2 learning. These findings have consistently suggested that executive attention, a component of WM, plays a pivotal role in the recognition of corrective feedback. Yet, limited attention has been devoted to examining the influence of WM in the context of L2 speech instruction based on task-oriented approaches. The present study extends our understanding by demonstrating that WM can exert a substantial impact on L2 speech acquisition when task-based instructional methods are employed.

In brief, these findings indicate that individuals with stronger auditory aptitude, as assessed through their ability to discern formants and their working memory capacity, tend to experience more substantial benefits from task-based L2 speech perception instruction. This conclusion aligns with emerging evidence from both cross-sectional and longitudinal studies, which suggests that success in learning a second language for conversational purposes is influenced by both auditory processing skills and working memory. This underscores the significance of taking specific auditory abilities, particularly formant discrimination, into consideration when designing instruction for L2 speech perception. Adapting instruction to target and enhance formant discrimination skills could prove to be a valuable strategy for enhancing L2 speech perception outcomes. It highlights the potential benefits of tailoring instructional methods to match learners' individual aptitude profiles, ultimately optimizing their learning outcomes. Furthermore, these findings lend support to the notion that variations in auditory aptitude among individuals play a pivotal role in determining the effectiveness of instructional interventions in the context of L2 speech perception.

7. Conclusion

In the context of our investigation into the acquisition of English vowel contrasts /e/-/æ/ and /ɪ/-/iː/ by 70 Chinese EFL learners, this study delved into the impact of TBLT on L2 speech learning. Specifically, we aimed to discern how TBLT contributes to L2 speech perception and how these instructional gains relate to distinct language aptitudes, namely, auditory processing and working memory. Our statistical analyses revealed three central findings that significantly contribute to our understanding of this area.

First, our study demonstrated that a concise 30-minute session of task-based instruction led to substantial improvements in L2 vowel perception among Chinese EFL learners. Remarkably, these enhancements extended beyond the trained lexical items, showcasing the broad utility of TBLT in enhancing overall L2 speech perception for this group. Secondly, our investigation unveiled that learners who possessed exceptional formant discrimination abilities experienced more pronounced gains in L2 speech perception following the instructional intervention. This underscores the pivotal role of this specific aptitude in facilitating L2 speech learning, especially when learners engage with TBLT. Third, our study observed that these instructional gains were most prominent among individuals with high levels of both formant discrimination ability and working memory capacity. This finding underscores the importance of considering learners' cognitive abilities when designing and implementing instructional approaches.

These results align with prior research indicating the effectiveness of task-based instruction in enhancing the perception of L2 speech features (Mora & Levkina, 2017; Gurzynski-Weiss et al., 2017; Loewen & Isbell, 2017; Solon et al., 2017) and the influence of individual aptitude factors on instructional effectiveness (Kissling, 2013). Furthermore, they underscore the echoing similar findings in the existing literature. Ultimately, our study contributes valuable pedagogical insights by highlighting the efficacy of TBLT in promoting L2 segmental perception accuracy and advocating for the thoughtful consideration of learners' cognitive abilities when designing instructional strategies in communicative-focused contexts. This research takes a further step toward shedding light on the intricate interplay between task-based instructional effectiveness and learners' individual aptitude factors.

8. Limitations and Future Directions

While this study has made significant strides in examining aptitude effects in meaning-oriented L2 phonetic training through Task-Based Instruction, it is essential to recognize the limitations encountered and chart potential future directions to refine our understanding of the mechanisms underlying successful instructed L2 speech learning.

One limitation of this study is the relatively small number of participants involved in the intervention, coupled with the relatively brief intervention period (30 minutes). Future research could consider expanding the participant pool and extending the intervention duration to investigate potential variations in outcomes, especially in terms of fluency. Longitudinal and controlled classroom-based studies may provide valuable insights in this regard.

The exclusive focus on segmental aspects of L2 phonology within this study represents another limitation. To offer a more comprehensive view of the aptitude-acquisition link, future research should incorporate a wider array of outcome measures, including controlled and spontaneous production, as well as the ability to decode prosodic emphasis at a sentence level (i.e., prosodic proficiency).

This research primarily examined the influence of instructional interventions on perceptual abilities in the context of second language acquisition. To obtain a more comprehensive understanding of the pedagogical advantages in L2 perception, forthcoming investigations should incorporate a range of perceptual assessments within experimental phonetics. These evaluations might encompass tasks involving the identification and differentiation of sounds using synthetic continua, evaluations conducted under conditions with background noise, and the potential inclusion of reaction time measures, as suggested by Iverson et al. (2003) and Lively et al. (1994).

To provide a more holistic understanding of the effectiveness of TBLT, future research should expand its focus beyond segmental accuracy. This entails investigating the influence of TBLT on L2 speech production, including various aspects of suprasegmental features, and evaluating both perception and production gains.

Future studies should aim to compare the effectiveness of TBLT to explicit instruction by incorporating a comparison group exposed to the same content through explicit instruction. This comparative approach will enable researchers to gain deeper insights into the nuances of instructional methods. This research involved a post-test administered one week after instruction, which, as per Spada and Tomita's (2010) synthesis of studies on instructed L2 acquisition, could be seen as a delayed rather than an immediate post-test.

To comprehensively gauge the long-term effectiveness of instruction, future research should contemplate adding a second post-test. This additional assessment would provide insights into the sustainability of gains attained through TBLT. Furthermore, upcoming studies should consider contextual variables, such as the research setting, whether it's a controlled laboratory environment or a classroom. It's worth noting that laboratory-based studies may yield more pronounced effects due to heightened experimental control. For instance, Li (2010) observed that the average impact of corrective feedback in laboratory-based studies ($d = 1.08$) was over twice that of classroom-based studies ($d = 0.50$).

By actively acknowledging and working to overcome these identified limitations, researchers have the potential to make substantial contributions towards achieving a deeper and more intricate comprehension of the processes and results associated with instructed second language (L2) speech learning within communicative-oriented environments. This proactive approach can lead to a richer understanding of how individuals acquire and use a second language for effective communication, shedding light on the complexities and nuances of this multifaceted educational endeavor.

Reference

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10/gfz97m

Atkins, P., & Baddeley, A. (1998). Working memory and distributed vocabulary learning. *Applied Psycholinguistics, 19*(4), 537-552. doi:10.1017/S0142716400010353

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? T*rends in Cognitive Sciences, 4,* 417–423.

Baills, F., Zhang, Y., Cheng, Y., Bu, Y., & Prieto, P. (2021). Listening to songs and singing benefitted initial stages of second language pronunciation but not recall of word meaning. *Language Learning, 71*(2), 369-413.

Best, C. T. (1995). Learning to perceive the sound pattern of English. *Advances in infancy research, 9,* 217-217.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language Learning & Language Teaching (Vol. 17, pp. 13–34).* John Benjamins Publishing Company. https://doi.org/10.1075/lllt.17.07bes

Bidelman, G. M., Gandour, J. T., & 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.

Brown, A. (1988). Functional Load and the Teaching of Pronunciation. *TESOL Quarterly, 22*(4), 593. https://doi.org/10.2307/3587258

Carroll, J. B., & Sapon, S. M. (1959). *Modern Language Aptitude Test (MLAT).* New York: Psychology Corporation.

Celce-Murcia M, Brinton DM, Goodwin JM, et al. (2010) *Teaching Pronunciation: A Reference for Teachers of English to Speakers of Other Languages. 2nd ed.* Cambridge: Cambridge University Press.

Chan, A. Y. W. (2010). Advanced Cantonese ESL learners' production of English speech sounds: Problems and strategies. System, 38(2), 316–328. https://doi.org/10.1016/j.*system.*2009.11.008

Chandrasekaran, B., Kraus, N., & Wong, P. C. (2012). Human inferior colliculus activity relates to individual differences in spoken language learning. *Journal of neurophysiology, 107*(5), 1325-1336.

Chandrasekaran, B., Sampath, P., & Wong, P. (2010). Individual variability in cue-weighting and lexical tone learning. *Journal of the Acoustical Society of America, 128* (1), 456–465. https://doi.org/10.1121/1.3445785

Darcy, I., Park, H., & 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/10.1016/j.lindif.2015.04.005

De Ridder, I., Vangehuchten, L., & Gomez, M. S. (2007). Enhancing Automaticity Through Task-based Language Learning. *Applied Linguistics, 28*(2), 309–315. https://doi.org/10.1093/applin/aml057

Dean, P. (2020, March 24). Digit Span Task (Visual & Text Entry). Gorilla. https://app.gorilla.sc/task/4698263

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching, 42*(4), 476–490. https://doi.org/10.1017/S026144480800551X

Doughty C, Williams J (1998) Pedagogical choices in focus on form. In: Doughty C, Williams J (eds) *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press, 197–285.

Ellis R, Shintani N (2014) *Exploring Language Pedagogy through Second Language Acquisition Research.* London: Routledge.

Ellis, N. C., & Sinclair, S. (1996). Working Memory in the Acquisition of Vocabulary and Syntax: Putting Language in Good Order. *Quarterly Journal of Experimental Psychology Section A, 49,* 234-250.

Ellis, R. (2001), Introduction: Investigating Form-Focused Instruction. *Language Learning, 51*: 1-46. https://doi.org/10.1111/j.1467-1770.2001.tb00013.x

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford, UK: Oxford University Press.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings: Task-based language teaching. *International Journal of Applied Linguistics, 19*(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Ellis, R., Baskturkmen, H., & Loewen, S. (2001). Learning uptake in communicative ESL lessons. *Language Learning, 51,* 281-318.

Field, J. (2005). Intelligibility and the Listener: The Role of Lexical Stress. *TESOL Quarterly, 39*(3), 399. https://doi.org/10.2307/3588487

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25,* 437–470. (n.d.).

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics, 25,* 437–470

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese Adults can Learn to Produce English /I/ and /l/ Accurately. *Language and Speech, 38*(1), 25–55. https://doi.org/10.1177/002383099503800102

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,* 3125–3134. https://doi.org/10.1121/1.413041

Fontanini, I., & Tomitch, L. M. B. (2009). Working memory capacity and L2 university students' comprehension of linear texts and hypertexts. *International Journal of English Studies, 9*, 1–18.

FRENCH, L., & O'BRIEN, I. (2008). Phonological memory and children's second language grammar learning. *Applied Psycholinguistics, 29*(3), 463-487. doi:10.1017/S0142716408080211

Gathercole, S. E. and E. V. Masoura (2005). "Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language." *Memory 13*(3-4): 422-429.

Giuliano, R. J., Pfordresher, P. Q., Stanley, E. M., Narayana, S., & Wicha, N. Y. (2011). Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change. *Frontiers in Psychology, 2,* 146.CrossRefGoogle ScholarPubMed

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612. https://doi.org/10.1146/annurev.psych.49.1.585

Goo, J. (2012). Corrective feedback and working memory capacity in interaction-driven L2 learning. *Studies in second language acquisition, 34*(3), 445-474.

Gordon, J. (2021). Pronunciation and Task-Based Instruction: Effects of a Classroom Intervention. *RELC Journal, 52*(1), 94–109. https://doi.org/10.1177/0033688220986919

Goswami, U. (2015). Sensory theories of developmental dyslexia: Three challenges for research. *Nature Reviews Neuroscience, 16*(1), 43–54. https://doi.org/10.1038/nrn3836

Gottfried TL, Miller JL, Payton PE. Effect of speaking rate on the perception of vowels. *Phonetica. 1990;47*(3-4):155-72. doi: 10.1159/000261860. PMID: 2151828.

Granena, G. (2013). Cognitive aptitudes for second language learning and the LLAMA language aptitude test. In G. Gisela & M. Long (Eds.), *Sensitive periods, language aptitude, and ultimate attainment* (pp. 105–130). Amsterdam, The Netherlands: John Benjamins.

Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in FTF and CMC modes. *Studies in Second Language Acquisition, 36,* 1-37.

Gurzynski-Weiss, L., Long, A. Y., & Solon, M. (2017). TBLT AND L2 PRONUNCIATION: DO THE BENEFITS OF TASKS EXTEND BEYOND GRAMMAR AND LEXIS? *Studies in Second Language Acquisition, 39*(2), 213–224. https://doi.org/10.1017/S0272263117000080

Hahn, L. D. (2004). Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals. *TESOL Quarterly, 38*(2), 201. https://doi.org/10.2307/3588378

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition, 14,* 25–38.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. J*ournal of the Acoustical Society of America, 108*(6), 3013–3022. https://doi.org/10.1121/1.1323463

Isaacs, T., Trofimovich, P., & Foote, J. A. (2018). Developing a user-oriented second language comprehensibility scale for English-medium universities. *Language Testing, 35*(2), 193–216. https://doi.org/10.1177/0265532217703433

Iverson P, Kuhl PK, Akahane-Yamada R, Diesch E, Tohkura Y, Kettermann A, Siebert C. A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition. 2003 Feb;87*(1):B47-57. doi: 10.1016/s0010-0277(02)00198-1. PMID: 12499111.

Jasmin, K., Sun, H., & Tierney, A. T. (2021). Effects of language experience on domain-general perceptual strategies. *Cognition, 206, Article 104481.* https://doi.org/10.1016/j.cognition.2020.104481

Jenkins, J. (2000). *The phonology of English as an international language.* Oxford: Oxford University.

Jung, Y., Kim, Y., & Murphy, J. (2017). THE ROLE OF TASK REPETITION IN LEARNING WORD-STRESS PATTERNS THROUGH AUDITORY PRIMING TASKS. *Studies in Second Language Acquisition, 39*(2), 319–346. https://doi.org/10.1017/S0272263117000031

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

Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research, 12*(2), 211–234. https://doi.org/10.1177/1362168807086288

Kissling, E. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal, 97*(3), 720–744. https://doi.org/ 10.1111/j.1540-4781.2013.12029.x

Ladefoged, P., & Johnson, K. (2014). *A course in phonetics.* Cengage learning.

Lee, A. H. & Lyster, R. (2016), Effects of different types of corrective feedback on receptive skills in a second language: A speech perception training study. *Language Learning, 66,* 809-833. https://doi.org/10.1111/lang.12167

Lee, B., Plonsky, L., & Saito, K. (2020). The effects of perception- vs. Production-based pronunciation instruction. *System, 88,* 102185.

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.

Levis, J. M. (2005). Changing Contexts and Shifting Paradigms in Pronunciation Teaching. *TESOL Quarterly, 39*(3), 369. https://doi.org/10.2307/3588485

Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning, 60,* 309-365. https://doi.org/10.1111/j.1467-9922.2010.00561.x

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, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F., & Doughty, C. J. (2013). Hi-LAB: A New Measure of Aptitude for High-Level Language Proficiency: High-Level Aptitude. *Language Learning, 63*(3), 530–566. https://doi.org/10.1111/lang.12011

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. I., & Yamada, T. (1994). Training Japanese listeners to identify English/r/and/l/. III. Long-term retention of new phonetic categories. *The Journal of the acoustical society of America, 96*(4), 2076-2087.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27*, 361-386.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in second language acquisition, 27*(3), 361-386.

Loewen, S., & Isbell, D. R. (2017). PRONUNCIATION IN FACE-TO-FACE AND AUDIO-ONLY SYNCHRONOUS COMPUTER-MEDIATED LEARNER INTERACTIONS. *Studies in Second Language Acquisition, 39*(2), 225–256. https://doi.org/10.1017/S0272263116000449

Long, M. H. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and assessing second language acquisition (pp. 77–99).* Clevedon, UK: Multilingual Matters.

Long, M. H. (1998). Focus on form in task-based language teaching. *University of Hawai'i Working Papers in ESL, 16,* 35–49.

Long, M. H. (2015). *Second language acquisition and task-based language teaching.* Malden, MA: Wiley.

Long, M. H., & Norris, J. M. (2000). *Task-based teaching and assessment. In M. Byran (Ed.), Encyclopedia of language teaching* (pp. 597–603). London, UK: Routledge.

Loschky, Lester & Bley-Vroman, Robert. (1993). Grammar and task-based methodology. *Tasks and Language Learning. 1.* 123-167.

Mackey, A., & Oliver, R. (2002). Interactional feedback and children's L2 development. *System, 30*(4), 459-477.

Mackey, A., & Sachs, R. (2012). Older learners in SLA research: A first look at working memory, feedback, and L2 development. *Language Learning, 62*(3), 704-740.

Mackey, A., Adams, R., Stafford, C., & Winke, P. (2010). Exploring the relationship between modified output and working memory capacity. *Language learning, 60*(3), 501-533.

McKinnon, S. (2017). TBLT INSTRUCTIONAL EFFECTS ON TONAL ALIGNMENT AND PITCH RANGE IN L2 SPANISH IMPERATIVES VERSUS DECLARATIVES. *Studies in Second Language Acquisition, 39*(2), 287–317. https://doi.org/10.1017/S0272263116000267

Mora, J. C., & Levkina, M. (2017). Task-based pronunciation teaching and research: Key issues and future directions. *Studies in Second Language Acquisition, 39*(2), 381–399. https://doi.org/10.1017/S0272263117000183

Mora-Plaza, I. (2021). What can I do when I travel? In L. Gurzynski-Weiss, & IATBLT (n.d.). The TBLT Language Learning Task Bank.

Mora-Plaza, I. Mora, J. C., & Gilabert, R. (2018). Learning L2 pronunciation through communicative tasks. In J. Levis (Ed.), *Proceedings of the 9th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, University of Utah, September, 2017 (pp. 174-184). Ames, IA: Iowa State University.

Mueller, J. L., Friederici, A. D., & 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

Munro, M. J., & Derwing, T. M. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System, 34*(4), 520–531. https://doi.org/10.1016/j.system.2006.09.004

Munro, M. J., & Derwing, T. M. (2008). Segmental Acquisition in Adult ESL Learners: A Longitudinal Study of Vowel Production. *Language Learning, 58*(3), 479–502. https://doi.org/10.1111/j.1467-9922.2008.00448.x

Munro, M.J., Derwing, T.M., & Thomson, R.I. (2015). Setting segmental priorities for English learners: Evidence from a longitudinal study. *International Review of Applied Linguistics in Language Teaching, 53,* 39–60.

Nation, 2012, The BNC/COCA word family lists, Victoria University of Wellington (2012)

Nuevo, A.-M., Adams, R., & Ross-Feldman, L. (2011). Chapter 7. Task complexity, modified output, and L2 development in learner–learner interaction. In P. Robinson (Ed.), *Task-Based Language Teaching (Vol. 2, pp. 175–202).* John Benjamins Publishing Company. https://doi.org/10.1075/tblt.2.13ch7

O'Brien, I., Segalowitz, N., Collentine, J., & Freed, B. (2006). Phonological memory and lexical, narrative, and grammatical skills in second-language oral production by adult learners. *Applied Psycholinguistics, 27*, 377–402.

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 48A(1),* 98–107. https://doi.org/10.1080/14640749508401378

Parlak, Ö., & Ziegler, N. (2017). THE IMPACT OF RECASTS ON THE DEVELOPMENT OF PRIMARY STRESS IN A SYNCHRONOUS COMPUTER-MEDIATED ENVIRONMENT. *Studies in Second Language Acquisition, 39*(2), 257–285. https://doi.org/10.1017/S0272263116000310

Perrachione, T. K., Lee, J., Ha, L. Y., & Wong, P. (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.

Povel, D. J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*(4), 411–440. https://doi.org/10.2307/40285311

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

Révész, A. (2014). Towards a fuller assessment of cognitive models of task-based learning: Investigating task-generated cognitive demands and processes. *Applied Linguistics, 35*, 87–92.

Revesz, A., & Han, Z. (2006). Task Content Familiarity, Task Type and Efficacy of Recasts. *Language Awareness - LANG AWARE, 15,* 160–179. https://doi.org/10.2167/la401.0

Révész, A., Sachs, R., & Mackey, A. (2011). *Task complexity, uptake of recasts, and second language development. In P. Robinson (Ed.), Researching task complexity: Task demands, task-based language learning and performance (pp. 203–238).* Amsterdam, the Netherlands: Benjamins.

Robinson, Peter. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies. 21*. 45-105.

Robinson, P. (2011). Task-based language learning: A review of issues. *Language Learning, 61* (Suppl. 1), 1–36.

Ruan, Y., & Saito, K. (2023). Less precise auditory processing limits instructed L2 speech learning: Communicative focus on phonetic form revisited. *System, 114,* 103020. https://doi.org/10.1016/j.system.2023.103020

Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *Tesol Quarterly, 46*(4), 842–854.

Saito, K. (2013). REEXAMINING EFFECTS OF FORM-FOCUSED INSTRUCTION ON L2 PRONUNCIATION DEVELOPMENT: The Role of Explicit Phonetic Information. *Studies in Second Language Acquisition, 35*(1), 1–29. https://doi.org/10.1017/S0272263112000666

Saito, K. (2019). To What Extent Does Long-Term Foreign Language Education Help Improve Spoken Second Language Lexical Proficiency? *TESOL Quarterly, 53*(1), 82–107. https://doi.org/10.1002/tesq.468

Saito, K. (2021). *Effects of Corrective Feedback on Second Language Pronunciation Development. In H. Nassaji & E. Kartchava (Eds.), The Cambridge Handbook of Corrective Feedback in Second Language Learning and Teaching (1st ed., pp. 407–428).* Cambridge University Press. https://doi.org/10.1017/9781108589789.020

Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing Auditory Precision Hypothesis-L2. *Language Teaching,* 1–17. https://doi.org/10.1017/S0261444822000453

Saito, K., & Brajot, F. X. (2013). Scrutinizing the role of length of residence and age of acquisition in the interlanguage pronunciation development of

English/by late Japanese bilinguals. *Bilingualism, 16*(4), 847-863. https://doi.org/10.1017/S1366728912000703

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of/ɹ/by Japanese learners of English. *Language Learning, 62*(2), 595–633.

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, 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., Hanzawa, K., Petrova, K., Kachlicka, M., Suzukida, Y., & Tierney, A. (2022). Incidental and Multimodal High Variability Phonetic Training: Potential, Limits, and Future Directions. *Language Learning, 72*(4), 1049–1091. https://doi.org/10.1111/lang.12503

Saito, K., Sun, H., & Tierney, A. (2020). Domain-general auditory processing determines success in second language pronunciation learning in adulthood: A longitudinal study. *Applied Psycholinguistics, 41*(5), 1083–1112. https://doi.org/10.1017/S0142716420000491

Saito, K., Suzukida, Y., Tran, M., & Tierney, A. (2021). Domain-general auditory processing partially explains second language speech learning in classroom set- tings: A review and generalization study. *Language Learning.*

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/10.1111/lang.12447

Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics, 37*(2), 217–240. https://doi.org/10.1017/S0142716414000502

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.

Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics, 39*(1), 187–224. https://doi.org/10.1017/S0142716417000418

Schmidt R (2001) *Attention. In: Robinson P (ed) Cognition and Second Language Instruction.* Cambridge: Cambridge University Press, 3–32.

Shao, Y., Saito, K., & Tierney, A. (2023). How Does Having a Good Ear Promote Instructed Second Language Pronunciation Development? Roles of Domain-General Auditory Processing in Choral Repetition Training. *TESOL Quarterly, 57*(1), 33–63. https://doi.org/10.1002/tesq.3120

Skehan, P. (1998). Task-based instruction. *Annual Review of Applied Linguistics, 18,* 268–286.

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). Amsterdam: John Benjamins.

Skehan, P. (2019). Language Aptitude Implicates Language and Cognitive Skills 1. In Z. (Edward) Wen, P. Skehan, A. Biedroń, S. Li, & R. L. Sparks (Eds.), *Language Aptitude (1st ed., pp. 56–77).* Routledge. https://doi.org/10.4324/9781315122021-4

Snowling MJ, Gooch D, McArthur G, Hulme C. Language Skills, but Not Frequency Discrimination, Predict Reading Skills in Children at Risk of Dyslexia. *Psychol Sci. 2018 Aug;29*(8):1270-1282. doi: 10.1177/0956797618763090. Epub 2018 May 23. PMID: 29791271; PMCID: PMC6088504.

Solon, M., Long, A. Y., & Gurzynski-Weiss, L. (2017). TASK COMPLEXITY, LANGUAGE-RELATED EPISODES, AND PRODUCTION OF L2 SPANISH VOWELS. *Studies in Second Language Acquisition, 39*(2), 347–380. https://doi.org/10.1017/S0272263116000425

Spada, N. (2011). Beyond form-focused instruction: Reflections on past, present and future research. *Language Teaching, 44*(2), 225–236. https://doi.org/10.1017/S0261444810000224

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language learning, 60*(2), 263-308.

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. https://doi.org/10.1017/S0272263120000649

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,* 1–23,

Suzukida, Y., & Saito, K. (2021). Which segmental features matter for successful L2 comprehensibility? Revisiting and generalizing the pedagogical value of the functional load principle. *Language Teaching Research, 25*(3), 431–450. https://doi.org/10.1177/1362168819858246

Swain M, Lapkin S (2001) Focus on form through collaborative dialogue: Exploring task effects. In: Bygate M, Skehan P, and Swain M (eds) *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing.* New York: Longman, 99–118.

Swan, M., & Smith, B. (Eds.). (2001). *Learner English: A teacher's guide to interference and other problems (2nd ed).* Cambridge University Press.

Tierney, A., White-Schwoch, T., MacLean, J., & Kraus, N. (2017). Individual differences in rhythm skills: Links with neural consistency and linguistic ability. *Journal of Cognitive Neuroscience, 29*(5), 855–868.

Tonkyn, A. (2012). Measuring and perceiving changes in oral complexity, accuracy and fluency: Examining instructed learners' short-term gains. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 221-246). Amsterdam, the Netherlands: Benjamins.

Wang, X., & Munro, M. J. (2004). Computer-based training for learning English vowel contrasts. *System, 32*(4), 539-552.

Yalçın, Ş., & Spada, N. (2016). LANGUAGE APTITUDE AND GRAMMATICAL DIFFICULTY: An EFL Classroom-Based Study. *Studies in Second Language Acquisition, 38*(2), 239–263. https://doi.org/10.1017/S0272263115000509

Yilmaz, Y., & Granena, G. (2016). The role of cognitive aptitudes for explicit language learning in the relative effects of explicit and implicit feedback. *Bilingualism: Language and Cognition, 19*(1), 147–161. https://doi.org/10.1017/S136672891400090X

Yilmaz, Y., & Granena, G. (2019). Cognitive individual differences as predictors of improvement and awareness under implicit and explicit feedback conditions. *The Modern Language Journal, 103*(3), 686-702.

Appendices

Appendix A. Questionnaire

Questionnaire

The aim of this questionnaire is to help us better understand your basic information and prior learning experience. Your honest and detailed responses will be greatly appreciated. Thank you very much!

1) Name: __________

2) Date of birth: __________

3) Gender: ___________

4) Your first language: ___________

5) **When did you start learning English? _____** years old

6) **How would you rate your proficiency in English?** (1-10 scale, with 1 being beginner and 10 being native speaker level) **_____**

7) **Did you receive any pronunciation instruction/training (taught by native or non-native teachers)?** If so what type of instruction/training? (e.g. Is the instruction on specific sounds, word stress or intonation? How long and how often? What materials do you use? What is the learning environment? etc.)

8) **Did you receive any music training?**

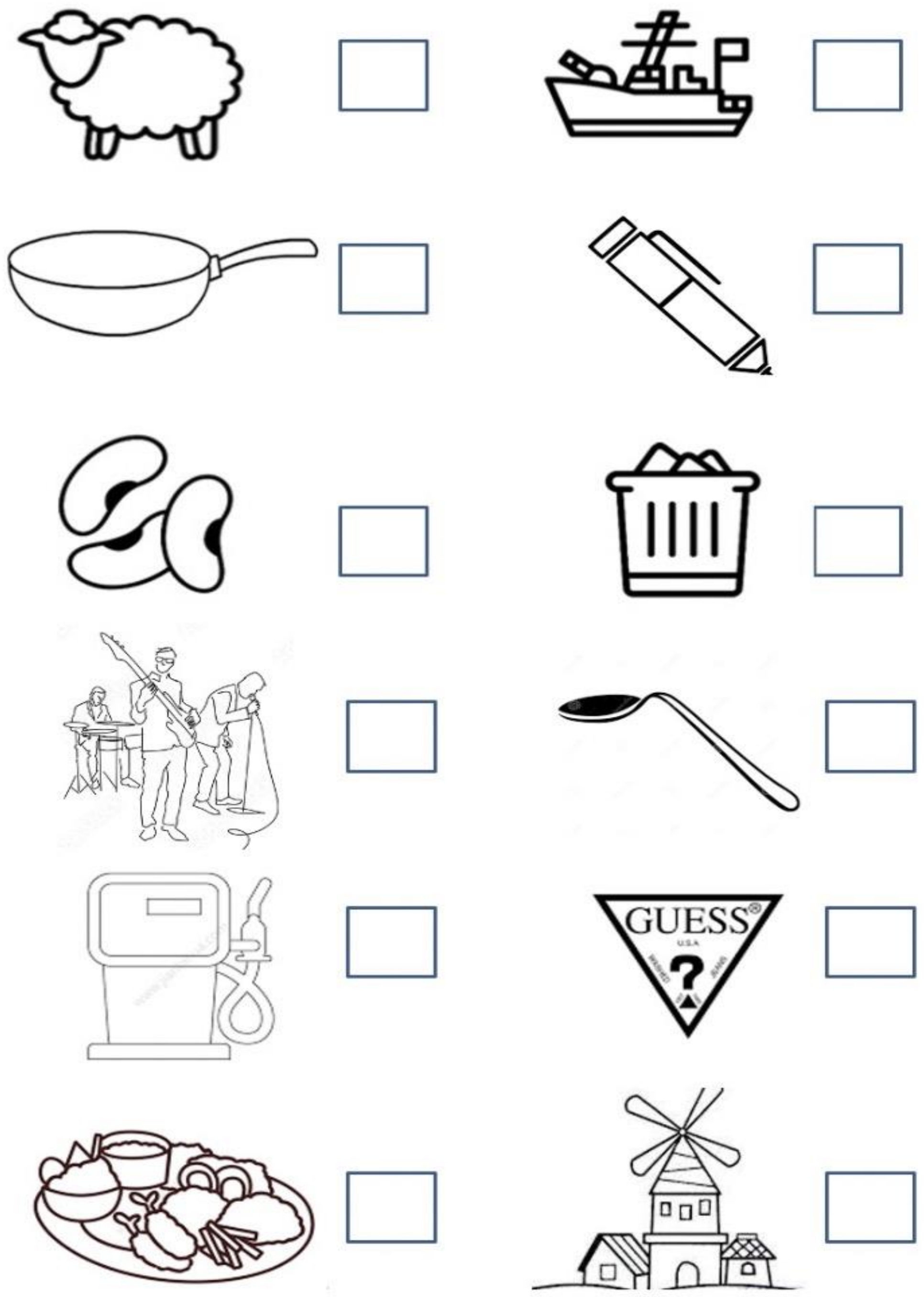
If so what type of training? (e.g. Is it vocal or instrumental? How long and how often? Is it for beginner, intermediate or advanced level? What is the learning environment? etc.)

Appendix B. Student Materials

STUDENT MATERIAL

Pre-task

1. Listen to your teacher and mark the correct picture with a cross.



2. Listen to your teacher and answer the following questions:

-Where do Sarah and Jake go on holiday? ______________________

-What activity are they going to do there? ______________________

Task

You are Peter and Lily.

You want to choose the **best UNUSUAL destinations** and **activities** to do on your summer holidays. You have been to a travel agency and they have offered you different plans. You have different preferences and conditions so you need to talk and decide on **3 TOP places** where you both would like to go.

1. **Read** your preferences (these are the places you want to go).
2. Be careful! You may have **different preferences** and **conditions** (and your partner doesn't know about them).
3. Make sure that you and your partner are **talking about the same** (there may be some confusable words [e.g., sheep/ship]).
4. Talk to your partner and decide on **3 top places** to go (15 mins).
5. **Circle the 3 places** where you would like to go **together** in the map and write down the activity next to it.
6. Present your plan to the travel agency. **Justify** why you chose the three destinations and, most importantly, why you gave up the others.

STUDENT A
PREFERENCES:

	visit the **Sheep** Museum –Edinburgh		visit a famous **pan** collector's house – New York
	go to famous **Marry**'s Holiday Club – Sydney		visit the museum of **Gas** – Los Angeles
	make **bins** by yourself (DIY) – Bangkok		find out London's best **chips** – London
	visit the mysterious **still** pond –Seoul		a traditional German **meal** – Berlin
	go to *the Ocean's* ***End*** which offers stunning views of the Atlantic Ocean – Nova Scotia		enjoy a jazz **band** performance – Cape Town

CONDITIONS：

- You love animals and natural scenery.
- You are allergic to beans.
- You major in history in school.
- You like cooking and exploring different restaurants.
- You have been learning Saxophone since you were a child.
- You are into fashion recently.

STUDENT B
PREFERENCES:

	visit the **Ship** Museum – Edinburgh		visit a famous **pen** collector's house – New York
	go to famous **Merry**'s Holiday Club – Sydney		visit the museum of **GUESS** – Los Angeles
	make **beans** by yourself – Bangkok		find out London's best **cheap** eats – London
	visit the world's largest **steel** plant – Seoul		a traditional German **mill** – Berlin
	explore the old town **and** historical park – Nova Scotia		visit Geller's (who can **bend** spoons with mind) Museum – Cape Town

CONDITIONS：

-You dislike anything related to animals.

- You are on a budget.

- You are a big fan of military and heavy industry

- You will have a stomachache if you eat potatoes.

- You are interested in history.

- You want to experience something strange and novel.

4

WORLD MAP

Nova Scotia
Edinburgh
London
New York
Berlin
Seoul
Los Angeles
Bangkok
Cape Town
Sydney

Appendix C. Teaching Guide

TEACHER'S GUIDE

Pre-task phase

1. Introduce the topic of "the UNUSUAL travelling plan" by showing photos related to this topic.

2. Elicit topic-related words and phrases from the students and add some target forms

3. First, read one of the words in the following minimal pairs twice and ask them to make a cross under the picture the word refers to. Then read the sentences and stress the word in bold once. Tell learners to check their answers.

sheep/ship,	*- A **ship** has always been my first option when I travel.*
pan/pen,	*-**Pen** collecting can teach you a lot about history and economy.*
bean/bin,	*- In Dublin city trucks take away all the **bins**.*
band/bend,	*- I'm not sure if we have time to enjoy a jazz **band** performance.*
gas/guess,	*-Have you ever been to the **Gas** Station Museum?*
meal/mill,	*-Lunch is the most important **meal** in Germany.*

4. Read the following text out loud and ask students to answer the question: Where do Sarah and Jake go on holiday? What activity are they going to do there?

*-Sarah, would you like to go to Germany to have a traditional German **meal**?*

*-I'm sorry, I have already tried once and I don't like German **meal**.*

*-What about visiting the world's largest **sheep** museum?*

*-Do you mean the world's largest **ship** museum in Scotland?*

-I'm afraid not. I'd love to visit an animal-related museum.

*-OK, perhaps you'd enjoy the world's famous **Gas** museum in Los Angeles.*

*-I don't think so as I'm not really interested in the history of gas station. But I like fashion. How about the museum of **GUESS**?*

-That sounds great! Let's go there!

5. Tell students that, in the task phase, they are expected to perform dialogues which are similar to those previously performed by the teacher.

Task phase

1. Put students in pairs or small groups.

2. Ask students to read the instructions on their worksheet carefully.

***As a teacher, you can include more or less conditions, depending on the difficulty you want to give to the task.

3.Record students while they are doing the task.

4. Monitor learners' productions and encourage students to practice the pronunciation of the target words.

5. Once students have finished, emphasize the importance of planning the task outcome before presenting it.

6. Ask learners to present their traveling palns.