

## Sequence Makes a Difference: Effects of Differential Blended Instructions on English

### Word Stress Identification by Chinese Teenager Learners

Yijie Zhou <sup>a</sup>, Yiwen Tang <sup>a\*</sup>, Jiaqiang Zhu <sup>b\*</sup>, Xiaoxiang Chen <sup>a</sup>

<sup>a</sup> School of Foreign Languages, Hunan University, Changsha, China

<sup>b</sup> Research Centre for Language, Cognition, and Neuroscience, Department of Language Science and Technology, The Hong Kong Polytechnic University, Hong Kong SAR, China

#### \* Correspondence:

Yiwen Tang: School of Foreign Languages, Hunan University, Lushan Road, Yuelu District, Changsha, China, [yiwentang@hnu.edu.cn](mailto:yiwentang@hnu.edu.cn)

Jiaqiang Zhu: Department of Language Science and Technology, The Hong Kong Polytechnic University, Hong Kong SAR, China, [jiaqiang.zhu@polyu.edu.hk](mailto:jiaqiang.zhu@polyu.edu.hk)

**Author contributions:** YZ and YT conceived and designed the study. YZ collected the data. YZ and JZ analyzed the data. YZ, YT, JZ, and XC drafted the manuscript. All the authors revised and approved the manuscript. YT and JZ finalized it for submission as corresponding authors.

**Conflict of Interest:** The authors have declared that no competing interests existed at the time of publication.

**Funding Statement:** The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by grants from the Social Science Project of Hunan Province in China (22YBA036), the Fundamental Research Funds for China MOOCs for Foreign Studies (cmfs-2023-0014) and the Research Funds for Reforming Postgraduate Education in Hunan Province in China (2023JGYB066).

**Ethics Approval Statement:** This study involving human participants was reviewed and approved by the ethics review board at the School of Foreign Languages of Hunan University. Written informed consent was signed by each participant.

**Data Availability Statement:** The raw data supporting the conclusions of this article will be

made available by the authors, without undue reservation.

## **Sequence Makes a Difference: Effects of Differential Blended Instructions on English**

### **Word Stress Identification by Chinese Teenager Learners**

#### **Abstract**

Previous studies highlighted the challenges faced by non-native learners in acquiring native-like word stress of English. This study delved into the impact of two blended instructions with differential sequences [explicit and then implicit instruction (EI); implicit and then explicit instruction (IE)] on English word stress identification by Chinese teenager learners of English, aiming to identify the possibly better blended instruction sequence. It also explored whether there is potential correlation between the instructional effects and memory systems. Eighty Chinese high school students participated in this five-week training study, with forty receiving EI and an equal number receiving IE. Their perceptual accuracy and reaction time collected from the pre-test and the immediate and delayed post-tests were compared and analyzed using the mixed-effects models. There were significant improvements in word stress identification through blended instruction, with the group of IE displaying better persistence effects. Besides, word stress identification improvements by EI were positively correlated with declarative memory, while perceptual improvements by IE exhibited a positive correlation with procedural memory. This study exhibited that the sequence of blended instruction led to differential learning and generalization effects, and suggested that IE could be more suitable for the improvements of word stress identification by Chinese teenager learners. Our findings lent support to the effectiveness of blended instruction, and provided the empirical evidence for the prediction of Declarative/Procedural Model on non-native language learning. Pedagogical implications for both typical and atypical populations were discussed.

**Keywords:** blended instruction; English word stress; identification; teenagers; development

## **I Introduction**

In recent years, research on the second language (L2) phonology has increasingly shifted from segmental features to suprasegmental aspects such as intonation and stress, which are essential for L2 comprehensibility and intelligibility (Bu & Zhou, 2021; Chen et al., 2001; Field, 2005; Lewis & Deterding, 2018). Among these, word stress stands out as a particularly challenging area for learners from non-stress languages like Chinese, making it a suitable and meaningful target for instructional research. Although models of stress typology suggest that speakers of non-stress languages can readily identify word stress in an L2 (Altmann & Vogel, 2002; Altmann, 2006), most empirical studies have focused on advanced adult learners from various L1 backgrounds (e.g., Chinese: Chen, 2013; Japanese: Saito & Saito, 2017; Korean: Lin et al., 2014). Far less is known about how Chinese teenager learners with lower L2 proficiency perceive English word stress. In fact, another line of research indicates that Chinese learners of English often experience difficulties in perceiving English word stress (Lai et al., 2010; Wang, 2008). Such mixed findings require a more systematic investigation towards English word stress identification by Chinese learners.

Furthermore, it has been widely reported that the acquisition of L2 phonology involves not only the explicit explanation of rules but also the implicit exposure to the target sounds (Cai, 2021; Ellis, 2002). Different from exploring the facilitative role of either implicit instruction or explicit instruction, recent studies focus on the *blended* instruction, which highlights that implicit and explicit instructions can be seen as complementary activities (Fainman & Tokar,

2019). In this aspect, researchers attempt to identify the sequence of implicit instruction and explicit instruction. Ellis (2009) advocates explicit explanation of rules before implicit teaching; conversely, Silveira (2002) suggests that teachers encourage direct imitation and provide explicit explanations only when students encounter difficulties in imitation. Though these studies investigated the effects of blended instruction, there is little evidence about the learning effect when implicit instruction and explicit instruction are integrally employed. To this end, the current study probed into the effects of blended instruction on the identification of English word stress, aiming to evaluate the better sequence of implicit instruction and explicit instruction, namely, which instruction method goes first can lead to larger gains, for Chinese teenager learners of English.

### **1 The Role of Contextual Sentences in L2 Word Stress Identification**

Word stress is a crucial prosodic feature that significantly contributes to L2 speech intelligibility and comprehensibility (Field, 2005; Lewis & Deterding, 2018). Early research on word stress perception primarily focused on isolated words and examined learners' sensitivity to acoustic cues such as pitch, duration, and intensity (Chrabaszcz et al., 2014; Zhang et al., 2008). However, recent studies have increasingly emphasized the role of higher-level linguistic information in word stress perception, including grammatical category, semantic meaning, and syntactic structure embedded in contextual sentences (Lee et al., 2019; Ortega-Llebaria et al., 2013; Wayland et al., 2006; Yu & Andruski, 2010). This is particularly important for stress contrasts in English (e.g., noun-verb pairs), where grammatical categories can facilitate more accurate stress placement (Davis & Kelly, 1997; Lai et al., 2010).

For example, Wayland et al. (2006) investigated the influence of syllabic structure and

lexical category on English stress perception among native Thai learners of English. Participants listened to pairs of nonwords that differed only in stress placement (e.g., “I’d like a TOOkips” vs. “I’d like a tooKIPS”). The findings indicated a strong influence of the learners’ first language (L1), supporting similar observations across various language backgrounds (Choi et al., 2019; Qin et al., 2017). In particular, speakers of non-stress languages where tone functions as a lexical property (e.g., Mandarin Chinese, Cantonese, and Vietnamese; Chen et al., 2023) are less likely to generalize stress patterns based on syllabic structure or lexical category.

In another study, Ortega-Llebaria et al. (2013) examined L2 learners’ ability to perceive stress in Spanish declarative sentences. Participants were asked to identify stressed and unstressed syllables within contextual sentences. The results suggested that some learners rely not only on prosodic cues but also on syntactic information such as word order and grammatical cues to infer stress positions. In addition, Ortega-Llebaria and Colantoni (2014) further compared the perception of sentence stress among Mandarin-speaking and Spanish-speaking learners of English, alongside native English speakers. Their findings showed that despite the typological similarities between English and Spanish, Spanish-speaking learners were more likely to rely on syntactic rather than prosodic cues to identify stress position in English as they did in Spanish (Kim, 2020; Lee et al., 2019; Zubizarreta & Nava, 2011). In contrast, Mandarin-speaking learners performed comparably to native speakers, aligning with predictions of the Stress Typology Model (Altmann & Vogel, 2002; Altmann, 2006), which posits that speakers of non-stress languages can readily identify the stress pattern in an L2, possibly due to their heightened sensitivity to stress patterns (Lai et al., 2010; Wang, 2008).

However, a key limitation of prior research is its focus on highly proficient adult learners

of English with different L1 backgrounds (e.g., Chinese: Yu & Andruski, 2010; Spanish: Ortega-Llebaria et al., 2013; Korean: Lin et al., 2014). It remains unclear whether Chinese teenager learners with lower English proficiency are able to accurately perceive English word stress when both auditory cues and contextual linguistic information are available.

## **2 The Blended Instruction on Non-Native Language Learning**

While previous studies have consistently shown that explicit instruction is more effective than implicit instruction in the acquisition of L2 grammar (Norris & Ortega, 2000; Spada & Tomita, 2010), far less is known about their relative effectiveness in L2 phonology, and the findings are inconclusive. Some research suggests that explicit instruction enhances learners' awareness of stress rules and improves perceptual accuracy of English word stress, particularly when the noun-verb word pairs are governed by regular patterns (Chen & Tian, 2020; Schwab & Dellwo, 2022; Yenkimaleki & Van Heuven, 2016). In contrast, other studies argue that implicit instruction may be more beneficial for learners from non-stress language backgrounds (e.g., Chinese, Korean, or Japanese), as immersed exposure to target sounds supports the perceptual development of stress patterns (Chan & Leung, 2014, 2018; Chen et al., 2025; Lindemann et al., 2016; Parlak et al., 2025; Saito & Saito, 2017). Nevertheless, these studies mainly used isolated words in perception tasks, while in the instruction phase the words with carrier sentences were involved. In addition, some studies have primarily focused on auditory processing in task design by providing isolated target words for learners to identify word stress (Chrabaszcz et al., 2014; Couper, 2022; Sippel & Martin, 2023). These discrepancies may lead to an underestimation of learners' actual abilities in word stress identification (Grant & Seitz, 2000), as the acquisition of explicit rules was not fully assessed. Hence, to assess

learners' perceptual abilities more comprehensively, it is necessary to incorporate contextual sentences into perception tasks (Lai et al., 2010; Lee et al., 2019; Ortega-Llebaria et al., 2013; Yu & Andruski, 2010).

Moreover, explicit and implicit instructions often coexist in authentic classroom settings (Ellis, 2011), suggesting the necessity to explore the blended instruction to combine the benefits (Fainman & Tokar, 2019). For example, Cai (2021) observed that blended instruction was effective in enhancing Chinese adult learners' perception of English sentence stress. However, it remained controversial regarding the sequence of implicit and explicit instructions on language learning. Shanks (2003) stated that language learning could benefit from implicit instruction prior to explicit instruction, even though learners were unable to verbalize what they had learned. Likewise, Silveira (2002) proposed that a teacher could encourage learners to imitate the target sounds through repetition, with explanation or scaffolding provided only when they fail to imitate. In contrast, Ellis (2009) argued that teachers could offer a metalinguistic explanation of the target rules before any practice activities, and then provide correct pronunciation materials for students to imitate and learn. These controversies need to be resolved so as to identify a more advantageous sequence of implicit and explicit instructions in non-native language learning.

Besides, blended instruction combining explicit and implicit instructions has also been shown to yield better persistent effects than using a single approach (Cai, 2018, 2021; Fainman & Tokar, 2019). Zibin and Altaxhaineh (2018) found that blended instruction significantly improved clause-combining skills in Jordanian EFL learners, with the experimental group outperforming the control group in subskills such as verb morphology and subordination. These

multimodal inputs led to durable learning gains and reinforced long-term retention through repeated exposure and active application. The retention benefits of blended instruction can be enhanced by multimodal input, which engages both auditory and visual channels. According to Mayer's (2014) Cognitive Theory of Multimedia Learning, dual coding (e.g., audio + captions) creates stronger memory traces than unimodal input. Audiovisual input, such as captions, helps learners break down speech into words, facilitating both listening and reading comprehension (Durbahn et al., 2020; Peters & Muñoz, 2020; Pujadas & Muñoz, 2019). It is especially useful for low-proficiency learners as a compensatory mechanism (Rodgers & Webb, 2017). In line with this, Saito et al. (2022) demonstrated that incidental training with audiovisual cues consolidates long-term memory for L2 vowel production. These findings align with our study design, where written contextual sentences complement auditory input, providing richer cognitive cues, reducing cognitive overload, and enabling more efficient information processing (Bisson et al., 2014).

Cai (2021) also found that the explicit-implicit (EI) sequence facilitated long-term retention of sentence stress patterns, as the explicit phase presented the systematic explanations of sentence stress rules while the implicit phase provided target input to help learners establish connections between form and meaning (Chen et al., 2025; Ellis, 2009; VanPatten & Cadierno, 1993). These findings collectively suggest that the EI sequence is suitable for English sentence stress acquisition. In contrast, the IE sequence might align better with English word stress identification. Word stress typically follows fixed patterns in isolated words (Anttila et al., 2020; Bu & Zhou, 2021), and its regularity allows learners to internalize these patterns through exemplar input. Graaff (1997) observed that implicit learning is more effective for the simple,

rule-governed structures (e.g., word stress) before explicit instruction. Given these characteristics, English word stress is supposed to be more suitable for the investigation of the effects of blended instruction among teenager learners, as recruited in our study, which has relatively simpler language structure and requires fewer cognitive resources (Cai, 2021). We hence estimated the effects of the blended sequence, either EI (explicit and then implicit instruction) or IE (implicit and then explicit instruction), that could possibly maximize the gains of English word stress identification among Chinese teenager learners.

### **3 Declarative/Procedural Model and Language Learning**

Based on Ullman (2001, 2016, 2020), the DP (declarative/procedural) model posited that two memory systems, including declarative memory and procedural memory, play significant roles in language learning and use. Declarative memory is critical for learning, storing, and consolidating factual knowledge of language, while procedural memory underlies the learning and processing of sequences and rules in language. Declarative memory has been further suggested to store the mental lexicon and retrieve semantic knowledge (Eichenbaum, 2004), and the learning of speech-sound category, articulation, and speech perception and production resort to procedural memory (Ullman, 2020). In addition, there was a seesaw effect predicted by DP model, namely, learning or processing in one memory system might block or inhibit the other (Ullman, 2016). For instance, retrieving irregular forms (e.g., dug) via declarative memory can hinder rule-based processing (dig: -ed) in procedural memory (Ullman, 2001, 2016, 2020).

The DP model also explains the underlying reasons why explicit and implicit instructions yield disparate outcomes (Ullman, 2016). Explicit instruction engages declarative memory,

while implicit learning promotes procedural memory use (Ullman, 2016, 2020). Correspondingly, studies by Morgan-Short et al. (2012) and Ji et al. (2023) confirm that implicit instruction is more closely tied to procedural memory, while learning outcomes via explicit instruction showed more reliance on learning in declarative memory. Furthermore, declarative memory is linked to explicit knowledge and controlled processing, while procedural memory supports implicit knowledge and automatic processing (Ellis, 2009; Ullman & Lovelett, 2018). Given the implicit nature of L2 phonological rules, features like suprasegmental patterns are likely acquired through procedural memory (Saffran et al., 1999).

Although prior studies have shown the procedural memory's role in acquiring syntax and intonation (e.g., Brill-Schuetz & Morgan-Short, 2014; Faretta-Stutenberg & Morgan-Short, 2018; Ji et al., 2023), it remains unclear whether the instruction sequence (EI or IE) differentially affects learners' reliance on these two memory systems in word stress identification.

#### **4 This Study**

In a nutshell, to investigate the effects of EI and IE on English word stress identification for Chinese teenager learners, the present study develops three related research questions: 1) What are the differences of instruction effects of EI and IE on the identification of English word stress? 2) What are the differences of persistence (retention) effects of EI and IE on the identification of English word stress? 3) Which learner-level factors are associated with the differential effects of EI and IE? A longitudinal study, lasting around 5 weeks in total, was designed to seek the answers, with pre-test and post-tests separated by the training/instruction phase. The test materials consisted of 15 English noun-verb word pairs in which 10 pairs served

as the trained items and 5 pairs were the untrained items.

For the selection of the target words, the difference between stressed and unstressed syllables was usually examined on noun-verb word pairs, which were identical in spelling forms and differed only in terms of stress placement (noun: stress on the first syllable; verb: stress on the second syllable; Beckman, 1986; Choi et al., 2019). As described earlier, some evidence suggested that native Mandarin speakers had difficulty in perceiving L2 English stress contrasts in a native-like manner (Wang, 2008). Accordingly, the current study attempted to investigate if any beneficial effects of differential blended instructions on English word stress identification could be observed among Chinese high school students. Improvements in word stress identification via blended instruction could be indexed by the higher accuracy and reduced reaction time (RT) in post-test(s) than in pre-test. Based on the related literature as reviewed before (Cai, 2021; Ellis, 2009; Silveira, 2002), it was hypothesized that the EI sequence presented better instruction effects whereas the IE sequence could exhibit better persistence effects. Meanwhile, it was also supposed that improvements on word stress identification might be positively correlated with declarative memory in the EI group, but the IE group's performance could be positively correlated with procedural memory, as indicated in previous studies (Faretta-Stutenberg & Morgan-Short, 2018; Ji et al., 2023; Morgan-Short et al., 2012; Ullman, 2001, 2016, 2020).

## **II Method**

### **1 Participants**

A total of 80 right-handed native Chinese participants (40 males, 40 females) aged from 14

to 16 were recruited from a high school in Hunan Province, China. They all were first-year students without formal and external musical training, and had learned English for at least 6 years at the time of testing. The participants were randomly assigned into two closely-matched groups receiving different sequences of blended instruction: a group received EI [mean age of the participants ( $M$ ) = 15.1,  $SD$  = 0.71, 20 males and 20 females], while the other group received IE ( $M$  = 15.2,  $SD$  = 0.59, 20 males and 20 females).

Language history questionnaire (Li et al., 2020) was employed to investigate participants' language proficiency in their L1 and L2. Results revealed that there were no significant differences between the EI and IE groups in the cumulative years of learning English ( $p$  = .242). All participants were confirmed with no experience of traveling to or living in English-speaking countries. In addition, vocabulary size test (Nation & Beglar, 2007) was used to check their vocabulary size and Oxford quick placement test (Syndicate, 2001) for L2 proficiency. Results of independent samples  $t$ -tests did not show any significant differences on vocabulary size ( $p$  = .937) or L2 proficiency ( $p$  = .951) between these two groups. None of our participants reported any hearing or speech problems, and they had normal or corrected-to-normal vision.

## **2 Memory Assessment**

### **2.1 Declarative Memory: Recognition Memory Task**

A linguistic version of the Recognition Memory Task (RMT) was conducted to assess participants' declarative memory. The RMT was presented in a visual modality, but not in an auditory one, to avoid potential associations between learning and phonological processing (Arthur et al., 2021). English non-words were selected from materials by Misyak et al. (2010). Similar to the procedure in Reifegerste et al. (2021), the RMT consisted of two phases of

incidental encoding and recognition, adapted from Lukács et al. (2017). E-Prime 3.0 was used for stimulus presentation.

During the encoding phase, participants were presented with 5 Chinese characters (referred to as “real”) and 5 disyllabic nonwords (referred to as “unreal” below). Participants were asked to decide if the words were real or unreal without indication for the subsequent test about their memory of these words. Participants were instructed to place their index fingers on the marked keys (“V” - real and “N” - unreal) on the keyboard. Each trial lasted for 500 milliseconds (ms), and then a fixation cross appeared at the center of the screen for 1000 ms. If a response occurred during the stimulus presentation, the trial ended at 500 ms. Alternatively, if a response occurred after 500 ms, a fixation cross replaced the object until participants made their response. The ten stimulus words during the encoding phase were presented in a random order.

The recognition phase took place about 5 minutes after the encoding phase. During this phase, participants responded by determining whether the stimuli presented on the screen were appearing for the first time (“new” stimulus) or had been presented during the learning phase (“old” stimulus) among 20 random trials (10 Chinese characters and 10 English non-words). Correct responses were calculated as the hit rate, while errors were identified as the false alarm rate. The  $d'$  score [ $d' = z(\text{Hit Rate}) - z(\text{False Alarm Rate})$ ] was ultimately calculated to measure declarative memory based on the Signal Detection Theory (Macmillan & Creelman, 2004).

## **2.2 Procedural Memory: Serial Reaction Time Task**

The serial reaction time task originated by Nissen and Bullemer (1987) was utilized to assess procedural memory. Participants were seated in front of a computer in a quiet room, and

directed to put four fingers from their dominant hand on keys on the keyboard. Letters “f, v, n, j” appeared in both random cycles and sequence cycles, respectively. They were instructed to watch for the target letter and to hit the corresponding key as quickly as possible when the letter appeared. According to Karatekin et al. (2007), the final sequential block was used as a manipulation check, ensuring that any observed increase in reaction time between the sequential block before the random block and the random block itself is not solely attributable to fatigue.

The reaction time data of incorrect trials were removed before calculating the average reaction time; then, correct reaction time was collected to measure procedural memory (Arthur et al., 2021). The shorter correct RT indicates the greater ability of the procedural memory. All participants’ demographic information can be found in Table 1.

### **3 Stimuli**

Fifteen real noun-verb word pairs of English were selected for perception tests in line with the methodology of previous studies (e.g., Beckman, 1986; Chrabaszcz et al., 2014).

All words were disyllabic minimal stress pairs (10 pairs as trained and 5 pairs as untrained stimuli), with either a trochaic or iambic stress pattern (e.g., SUBject: noun; subJECT: verb). To retain naturalness, sound of each target word downloaded from the *Longman Dictionary of Contemporary English* was played in isolation in the identification task, and the relevant contextual sentence was presented on the screen. Since the goal of this experiment was to examine the effects of blended instruction on word stress identification, some steps were taken to investigate whether learners were fully aware of the word stress rules of disyllabic words. Similar to Lai et al. (2010) and Zhang et al. (2008), information regarding the lexical categories

of the target words was conveyed to participants by embedding the target words in relevant contextual sentences (see Table S1 in the supplementary materials for all stimuli of this study).

#### **4 Procedure**

The training/instruction was implemented between pre-test and post-tests, and this whole training study lasted about five weeks. All participants were first given a series of background tests to obtain a base for homogeneity, including the control tests (language history questionnaire, Oxford quick placement test, and vocabulary size test) and the memory tests (recognition memory task and serial reaction time task). The instructional materials for both groups were the same. Figure 1 illustrates the concrete experimental procedure and blended instruction applied in both groups.

**[insert Figure 1.]**

##### **4.1 The Identification Task**

In line with prior research (Cebrian et al., 2024; O'brien, 2019), this study employed an identification task where participants were required to listen to an auditory stimulus and select the corresponding sound. In addition, as learners were exposed to both explicit word stress rules and implicit input through blended instruction, multimodal cues (e.g., auditory stimuli and contextual sentences) were simultaneously incorporated into the identification task to assess improvements across the two groups. The identification task was conducted using the computer software of xPerception (Xiong, 2019) in quiet environment. The pre-test, immediate post-test, and delayed post-test shared identical stimuli and common procedures.

Participants were guided to carefully read the instructions displayed on the screen, and the experimenter provided a detailed explanation of the test procedures and requirements in the

participants' native language, to ensure that they understood the test protocols. All these 30 words from the word pairs were employed. For each trial, the sound was played in isolation with the relevant contextual sentences showing on the screen at the same time. As an example, after listening to a sound with contextual sentence in a trial, participants were required to decide which syllable was the stressed syllable ("A" referred to the first syllable, while "B" was the second syllable) within 5 seconds. They were encouraged to respond as quickly as possible without sacrificing the accuracy. The order of trials was randomly presented. Accuracy and reaction time were collected to measure participants' performance of identifying English word stress.

#### **4.2 The Instruction Phase**

Formal instructions were conducted twice a week (Wednesday and Friday mornings) for two weeks, and each class lasted 40 minutes. The teaching processes in these two groups were conducted by the same experimenter skillful in teaching. The first week involved conducting the pre-test, and the immediate post-test was launched after the two-week blended instruction. In the fifth week, a delayed post-test was conducted to examine any potential persistence effects, in line with Keck et al. (2006). In the EI group, participants received the EI sequence, comprising explicit instruction first and then implicit instruction. Participants in the IE group received the IE sequence, consisting of implicit instruction first and then the explicit instruction.

During the explicit phase, English word stress was taught through the systematic explanation of rules and the practice of using these rules. The acoustic features of stressed and unstressed syllables were compared explicitly, including pitch, intensity, and vowel duration

(Kim et al., 2025) by spectrograms to help students understand the word stress better. As exemplified in Table 2, it can be seen that the pitch and intensity of the stressed syllable in “CONduct” are higher than the unstressed syllable, and duration of stressed vowel is longer than that of the unstressed vowel. In contrast, for “conDUCT”, the second syllable had higher pitch and intensity and longer vowel duration than the first syllable. These manifestations are generally in line with acoustic information of English word stress, either for the trochaic or the iambic stress pattern (Choi et al., 2019).

Besides, the basic rules for word stress of noun-verb pairs were explained (i.e., nouns tend to have stress on the first syllable, but verbs tend to have stress on the second syllable). The concept and role of word stress in English were explained with examples of how word stress rules affected the pronunciation and meaning of words, and the relevant contextual sentences of target words were provided. In line with prior studies (Horgues, 2013; Ortega-Llebaria et al., 2013), we used simple declarative sentences which were less challenging and demanding for the understanding by teenager learners, and these sentences in training were different from the contextual sentences used in the identification task. Taking “contract” as an example. When the word stress was on the first syllable, it is a noun signifying a legal document that states and explains a formal agreement with contextual sentence (This is a “contract”). While the second syllable was stressed, it is a verb referring to being smaller in amount or in size with relevant contextual sentence (The metal “contracted”). In addition to systematic explanation, students in the EI group were encouraged to analyze rules with trained words beyond the examples taught by the teacher and to practice constructing simple sentences with similar structures.

In implicit phase, students were instructed in a natural and spontaneous learning

environment without any metalinguistic knowledge of English word stress. They were exposed to the recordings of 20 target words and tried to guess the lexical category or stress pattern of each trained word. Students were encouraged to determine the stressed and unstressed syllables of words while listening to the audios of target words, marking the stressed and unstressed syllables with “Oo” (the first syllable was stressed) and “oO” (the second syllable was stressed). Although learners received immersed input of target sounds, they were not introduced to any rules of word stress. Moreover, they were engaged in extra imitating activity, which aided them to notice the gap between their own pronunciations and sounds of target structures (Saito, 2019). Their imitations were corrected only by implicit feedback in which the teacher repeated words with correct stress patterns and provided positive reinforcement, aiming to encourage students to correct their errors and continue internalizing the correct pronunciations, rather than relying on direct rule-based guidance.

An important note, as suggested by one anonymous reviewer, involved that it was less likely to define a pure implicit or explicit learning; instead, implicit learning is relative to explicit learning. In this study, we followed previous studies to design our teaching methods, e.g., imitation with implicit feedback is implicit relative to direct explanation of stress rules (Fukuta & Yamashita, 2023; Robinson, 1997).

## **5 Data Analysis**

Accuracy and reaction time from a total of 7200 trials (80 participants × 3 test sessions × 30 words) were collected in the initial analysis. Accuracy in each trial was coded as 0 or 1 (incorrect or correct) for each participant. Of those, to reduce the impact of outliers on model parameter estimation, 0.8% of the accuracy (58) and 1.7% of RT (120) data were excluded

which were characterized by absolute standardized residuals greater than 2.5 (Baayen & Milin, 2010).

The statistical analysis was performed via a generalized linear mixed-effects model (GLMM) for accuracy and a linear mixed-effects model (LMM) for RT using the “lme4” package (Bates et al., 2020) in R 4.3.2 (R Core Team, 2023). Due to the right-skewed distribution of RT data, it was advisable to transform the raw RT data logarithmically when building LMM (Ma et al., 2022); but, for brevity, the raw RTs were reported and described in Results section. The models were constructed with Group (EI group, IE group), Type (trained, untrained), Session (pre-test, immediate post-test, delayed post-test) and their interactions acting as fixed effects with the dependent variable of accuracy in GLMM and RT in LMM. By-“Subject” and by-“Item” random slopes were prioritized unless there was a convergence issue or no significant difference was found when compared with simpler models with random intercepts through likelihood ratio tests. The ANOVA function in “lmerTest” package was employed to compare the initial model with a simplified model that excluded a specific fixed factor (Kuznetsova et al., 2017). Post-hoc pairwise comparisons were calculated by applying the function of R package “lsmeans” with Tukey adjustment (Lenth, 2016). Meanwhile, we also constructed multiple regression models to examine the relationship between our teenager learners’ identification performance and their declarative/procedural memory capacities, using the “MASS” package (Ripley et al., 2012) in R. The statistical significance level is 0.05 in this study.

### **III Results**

To illustrate the instruction effects and persistence effects of EI and IE, accuracy and RT data were compared and analyzed from the pretest, immediate post-test, and delayed post-test.

## 1 Accuracy

Figure 2 plots accuracy by the EI and IE groups across three test sessions on trained and untrained words. It displayed that accuracy of word stress identification was improved in both groups. In immediate post-test, accuracy in the EI group was similar to the IE group on trained words (EI:  $M = 0.831$ ,  $SD = 0.07$ ; IE:  $M = 0.833$ ,  $SD = 0.10$ ), but was lower on untrained words (EI:  $M = 0.793$ ,  $SD = 0.08$ ; IE:  $M = 0.843$ ,  $SD = 0.09$ ). While in delayed post-test, the IE group exhibited higher accuracy than the EI group on both trained (EI:  $M = 0.820$ ,  $SD = 0.07$ ; IE:  $M = 0.838$ ,  $SD = 0.09$ ) and untrained words (EI:  $M = 0.788$ ,  $SD = 0.11$ ; IE:  $M = 0.840$ ,  $SD = 0.09$ ).

**[insert Figure 2.]**

The GLMM for accuracy firstly revealed a significant main effect of session,  $\chi^2(2) = 247.690$ ,  $p < .001$ . There were also three interactions of Group  $\times$  Session,  $\chi^2(2) = 6.299$ ,  $p < .05$ , Group  $\times$  Type,  $\chi^2(1) = 4.269$ ,  $p < .05$ , and Session  $\times$  Type,  $\chi^2(2) = 9.656$ ,  $p < .01$ .

To further explore the interactions, post-hoc pairwise comparisons were conducted. The accuracy obtained by two groups in immediate post-test was statistically higher than pretest, EI:  $\beta = -2.017$ ,  $SE = 0.439$ , z.ratio ( $z$ ) =  $-4.59$ ,  $p < 0.001$ ; IE:  $\beta = -2.245$ ,  $SE = 0.441$ ,  $z = -5.09$ ,  $p < 0.001$ , indicating the effective instruction effects of two blended instructions. Meanwhile, in immediate post-test, there was no difference between the EI group and the IE group on trained and untrained words ( $\beta = -.210$ ,  $SE = 0.151$ ,  $z = -1.39$ ,  $p = 0.164$ ), showing the similar perceptual improvements within two groups.

In delayed post-test, accuracy within two groups did not differ from that of immediate

post-test on both trained and untrained words. However, accuracy obtained by the IE group was significantly higher than accuracy by the EI group on untrained words ( $\beta = -0.459$ ,  $SE = 0.155$ ,  $z = -2.96$ ,  $p < 0.01$ ), indicating the stronger retention effects on untrained words in the IE group.

To summarize, accuracy of English word stress identification was significantly increased through EI and IE. Results in immediate post-test implied the similar instruction effects of EI and IE on trained and untrained words. Besides, the persistence of learning gains on trained words within two groups remained comparable in delayed post-test but was stronger in the IE group.

## 2 Reaction Time

Figure 3 depicts RT by the EI group and the IE group across three test sessions, faceted by trained and untrained words with error bars illustrating 95% confidence intervals. Overall, this figure reflected a decreased tendency of RT in the tests. In immediate post-test, RT in the EI group was shorter than the IE group on trained (EI:  $M = 1717.300$ ,  $SD = 97.24$ ; IE:  $M = 1841.725$ ,  $SD = 233.26$ ) and untrained words (EI:  $M = 1738.7$ ,  $SD = 87.94$ ; IE:  $M = 1872$ ,  $SD = 190.91$ ). In delayed post-test, however, the IE group demonstrated shorter RT than the EI group on trained (EI:  $M = 1847.025$ ,  $SD = 188.17$ ; IE:  $M = 1766.801$ ,  $SD = 162.62$ ) and untrained words (EI:  $M = 1948.225$ ,  $SD = 207.89$ ; IE:  $M = 1663.930$ ,  $SD = 158.63$ ).

**[insert Figure 3.]**

First, the LMM uncovered two significant main effects of session [ $\chi^2(2) = 39159.632$ ,  $p < 0.001$ ] and type [ $\chi^2(1) = 7.959$ ,  $p = 0.005$ ]. There were three interactions of Group  $\times$  Session,  $\chi^2(2) = 1186.091$ ,  $p < 0.001$ , Group  $\times$  Type,  $\chi^2(1) = 70.025$ ,  $p < 0.001$ , and Session  $\times$

Type,  $\chi^2(2) = 33.879, p < 0.001$ , and, notably, a three-way interaction of Group  $\times$  Session  $\times$  Type,  $\chi^2(2) = 157.290, p < 0.001$ .

Further pairwise comparisons demonstrated that RTs were significantly shorter (indicating faster responses) in the immediate post-test than in the pre-test for both groups on trained items [EI:  $\beta = 0.410, SE = 0.004, t.\text{ratio} (t) = 107.85, p < 0.001$ ; IE:  $\beta = 0.363, SE = 0.004, t = 96.18, p < 0.001$ ] and untrained words (EI:  $\beta = 0.408, SE = 0.005, t = 75.39, p < 0.001$ ; IE:  $\beta = 0.336, SE = 0.005, t = 62.61, p < 0.001$ ), implying the efficient learning outcomes. Nevertheless, pertaining to the between-group difference, the EI group had shorter RT than the IE group on trained ( $\beta = -0.064, SE = 0.012, t = -5.28, p < 0.001$ ) and untrained ( $\beta = -0.076, SE = 0.013, t = -6.03, p < 0.001$ ) words.

Interestingly, in delayed post-test, RT increased in the EI group (trained:  $\beta = -0.076, SE = 0.004, t = -19.95, p < 0.001$ ; untrained:  $\beta = -0.104, SE = 0.005, t = -19.29, p < 0.001$ ), indicating slowed responses. Conversely, RT decreased in the IE group (trained:  $\beta = 0.038, SE = 0.004, t = 9.99, p < 0.001$ ; untrained:  $\beta = 0.122, SE = 0.006, t = 22.08, p < 0.001$ ), indicating faster responses. Contrary to the shorter RT in the EI group in immediate post-test, the IE group revealed shorter RT in delayed post-test, which implicated the more robust persistence and internalization of learning gains.

Collectively, analysis of RT revealed the beneficial effects of blended instruction. In immediate post-test, the more significantly decreased RT in EI reflected better instruction effects than IE. However, in delayed post-test, participants under IE responded to the word stress identification task faster than EI on trained and untrained words, suggesting the stronger persistence effects.

### 3 Declarative Memory and Procedural Memory

In order to further examine the factors associated with the varying effects of EI and IE, multiple linear regression models were established with the dependent variable of accuracy in immediate post-test and delayed post-test. Declarative memory and procedural memory were treated as independent variables to respectively examine their relationship with word stress identification in the EI group and the IE group.

The regression models for accuracy on trained and untrained words were statistically significant. Specifically, the Durbin-Watson values for each regression model fell between 1 and 3, suggesting the absence of autocorrelation in residuals (Jović, 2016). The Variance Inflation Factor values for declarative memory and procedural memory, ranging from 1 to 5, indicated that there was no linear correlation between these two types of memory. These results were supportive of the correlation between memory system and word stress identification in the current study (see the detailed results in Table S2 in the supplementary materials).

**[insert Figure 4.]**

As shown in Figures 4 and 5, significant correlations were found in both groups during immediate post-test. For trained words, accuracy in the EI group was negatively correlated with procedural memory ( $\beta = -0.058, t = -9.35, p < 0.001$ ) and positively correlated with declarative memory ( $\beta = 0.086, t = 2.85, p = 0.007$ ). For the IE group, however, accuracy was negatively correlated with declarative memory ( $\beta = -0.175, t = -2.41, p = 0.021$ ) but positively correlated with procedural memory ( $\beta = 0.062, t = 4.39, p < 0.001$ ). Moreover, for untrained words, accuracy of the EI group exhibited singularly negative correlation with procedural

memory ( $\beta = -0.079, t = -4.16, p < 0.001$ ). Accuracy of the IE group had significantly positive correlation with procedural memory ( $\beta = 0.057, t = 4.07, p < 0.001$ ) and marginally negative correlation with declarative memory ( $\beta = -0.129, t = -1.77, p = 0.084$ ).

**[insert Figure 5.]**

Similarly, in delayed post-test, significant correlations were observed in these two groups. For trained words, accuracy in the EI group, again, revealed the same correlation with declarative memory ( $\beta = 0.197, t = 5.05, p < 0.001$ ) and procedural memory ( $\beta = -0.032, t = -4.01, p < 0.001$ ). The accuracy of the IE group demonstrated significantly negative correlation with declarative memory ( $\beta = -0.123, t = -2.28, p = 0.028$ ) and positive correlation with procedural memory ( $\beta = 0.062, t = 5.95, p < 0.001$ ). For untrained words, the IE group displayed singularly positive correlation with procedural memory ( $\beta = 0.053, t = 2.97, p = 0.005$ ). In the EI group, accuracy was positively correlated with declarative memory ( $\beta = 0.160, t = 3.15, p = 0.003$ ) and negatively correlated with procedural memory ( $\beta = -0.074, t = -7.08, p < 0.001$ ). For more details, the above results can be found in Table S3 in the supplementary materials. In summary, the correlations between accuracy and memory systems indicated that the EI group relied predominantly on declarative memory, whereas the IE group depended more on procedural memory in word stress identification.

#### **IV Discussion**

The current study investigated the effects of EI and IE on the identification of English word stress so as to identify the potentially optimal blended sequence for word stress acquisition for Chinese teenager learners. Results of accuracy and RT were analyzed to provide

answers to three research questions: 1) What are the differences of instruction effects of EI and IE on the identification of English word stress? 2) What are the differences of persistence (retention) effects of EI and IE on the identification of English word stress? 3) Which learner-level factors are associated with the differential effects of EI and IE? We hypothesized that the EI group demonstrated better instruction effects, while the IE group exhibited better persistence effects in the identification of English word stress. Besides, we assumed that effects of these two blended sequences might be differently related with declarative memory and procedural memory abilities. The main findings were separately reported below.

### **1 The Instruction Effects of EI and IE on English Word Stress Identification**

The results of immediate post-test revealed that compared to the pre-test, there were significant improvements on (higher) perceptual accuracy and (shorter) reaction time on both trained and untrained words in the two groups, exhibiting the effectiveness of EI and IE on English word stress identification by Chinese teenager learners. Meanwhile, despite the similar accuracy, the EI group further showed shorter reaction time than the IE group on trained and untrained words, suggesting the better instruction and generalization effects of the EI sequence.

The EI group's shorter reaction times may reflect the cognitive benefits of early rule-based instruction, which serves as metalinguistic scaffolding during the learning process (Ellis, 2005; Roehr-Brackin, 2014). By receiving explicit instruction on English word stress rules, learners were able to develop metalinguistic knowledge early on (Chen & Tian, 2020; Schwab & Dellwo, 2022), which in turn facilitated more efficient and rapid decision-making by reducing the processing load (Roehr-Brackin, 2014). The gains in explicit knowledge of word stress likely enabled learners in the EI group to identify stress patterns more effectively.

Furthermore, the metalinguistic cues embedded in the identification task may have helped them apply the rules to distinguish the lexical category of the target words and determine stress assignment accordingly (Ortín & Simonet, 2023). In addition, the multimodal learning task, including auditory stimuli and visually presented contextual sentences (e.g., “He might refuse the offer”), provided semantic and grammatical cues that encouraged top-down predictive processing (Batel, 2020; Cope et al., 2017; Ortega-Llebaria et al., 2013; Ortega-Llebaria & Colantoni, 2014). These combined factors likely contributed to the EI group’s significantly shorter reaction time compared to the IE group in the immediate post-test. Nevertheless, while efficient, such rule-based processing may lead to a heavy reliance on metalinguistic knowledge rather than natural auditory processing (Ellis, 2002).

Despite comparable accuracy between these two groups on trained words, the IE group demonstrated better performance on untrained words, highlighting the advantage of implicit learning. In the IE sequence, learners were immersed in repeated exposure to exemplars, which enabled them to internalize the probabilistic regularities of word stress patterns without explicit rule instruction (Chan & Leung, 2014; Saito & Saito, 2017; Zhu et al., 2023a). Based on the Stress Typology Model (Altmann & Vogel, 2002), Chinese learners had heightened sensitivity to English word stress on account of the typological differences between Chinese (a non-stress language) and English (a stress language). Some studies suggest that Mandarin-speaking L2 learners’ processing of English word stress relied more on F0 (fundamental frequency) to identify the stressed and unstressed syllables (Chrabaszcz et al., 2014; Qin et al., 2017; Zhang et al., 2008), which might have further supported the bottom-up processing fostered by the IE sequence. This exemplar-based learning prompts the gradual internalization of probabilistic

regularities of word stress patterns (Chen, 2013; Ellis, 2002, 2005), which tends to be more robust in generalizing to novel items. Although implicit knowledge usually requires more time to consolidate (Ellis, 2009), it is readily available through automatic processing and unplanned language use, potentially explaining the IE group's higher accuracy on untrained words. In contrast, the EI group relied more heavily on explicit rule learning, which could result in their accurate performance on trained items but led to poorer generalization due to reduced internalization of target patterns (Cai, 2018).

## **2 The Persistence Effects of EI and IE on English Word Stress Identification**

Generally, as shown in delayed post-test, the improvements within the two groups were persistent, regardless of the trained and untrained words. This provided compelling evidence for the effectiveness of blended instruction on English word stress acquisition. It was noted that the IE group had higher accuracy on untrained words and shorter RT on both trained and untrained words. Our study hence offers one more enlightenment that the IE sequence exhibited better persistence effects than the EI sequence.

The different sequence between explicit and implicit instructions influenced the reliance on multimodal cues. The IE sequence encouraged learners to depend more on auditory processing while potentially reducing reliance on grammatical or semantic cues in sentences. This type of implicit learning mechanism, similar to that of native speakers, was more conducive to the internalization of word stress patterns (Chan & Leung, 2018; Parlak et al., 2025). In contrast, the EI sequence led learners to rely more on grammatical or semantic cues in contextual sentences to identify word category and, subsequently, stress placement. While the explicit learning of word stress yielded better immediate effects, it showed weaker long-term

retention.

Furthermore, previous studies have shown that lower-proficiency learners, due to limited linguistic knowledge, tend to rely on simpler, more intuitive linguistic units (e.g., grammatical or semantic cues) when identifying word stress, as these cues are easier to process and reduce cognitive load (Ingram, 2005; Wayland et al., 2006). However, the blended instruction in this study may facilitate these Chinese teenager learners with lower proficiency in identifying English word stress through multiple linguistic cues (Ellis, 2005; Cai, 2021). Contextual sentences, by providing more intuitive and less cognitively demanding grammatical or semantic cues, offer an effective way for learners to identify word stress. In the EI sequence, learners tended to focus on grammatical cues of target words in contextual sentences before shifting their attention to auditory processing (Nagle & Hiver, 2024). This allowed them to integrate both grammatical and auditory cues in word stress identification. In contrast, learners in the IE group might focus on the stress pattern first and subsequently verify their auditory hypotheses through contextual sentences (Ortega-Llebaria et al., 2013; Saito & Saito, 2017; Yu & Andruski, 2010). This approach aligns with N. C. Ellis' (2002) proviso to put foreign language learning in proper balance, which suggests that learners first register linguistic representations and then verify and reinforce their hypotheses by explicit rules. As a result, it may explain why the EI group exhibited longer reaction time than the IE group in the delayed post-test.

Besides, learners with lower language proficiency may rely solely on auditory cues or isolated grammatical cues (Cai, 2021; Frost, 2011). In the IE sequence, learners first internalize the stress patterns through repeated exposure to auditory input in a bottom-up way

(Chan & Leung, 2014, 2018; Jung et al., 2017), which are later confirmed or refined through explicit instruction. This could facilitate learners' extraction and retention of the target structure, thus enabling it to be stored in long-term memory (Ruchkin et al., 2003). As a result, the IE group performed better in integrating both the auditory and grammatical cues in the word stress identification. In contrast, the rule explanation perhaps reduces learners' curiosity and attentional scope during the subsequent implicit phase. The possession of explicit rules might lead to a sense of fatigue or diminished satisfaction, which results in less engagement and motivation to refine their learning of the target structure and potentially affects long-term learning outcomes (Moyer, 1999; Uchihara et al., 2022).

### **3 The Factors Associated with the Varying Effects on English Word Stress Identification**

Our results manifested that there was a strongly positive correlation with declarative memory and a negative correlation with procedural memory in the EI group. In contrast, English word stress acquisition correlated positively with procedural memory and negatively with declarative memory in the IE group. These findings validate the prediction of the DP model on speech perception, and provide novel insights that learning via the EI sequence primarily relies on the declarative memory system, while learning by the IE sequence mainly resorts to the procedural memory system.

According to DP model (Ullman, 2001), speech perception and production by native speakers largely depend on procedural memory, while L2 learners mainly rely on declarative memory (Ullman, 2020). Instructional methods could facilitate the storage and processing of novel information from working memory to long-term memory (Sweller et al., 2019), and different instructions (different instruction sequences in this study) may exert variable effects

on L2 learners' reliance on these two memory systems (Ullman, 2016). The EI group processed the target form based more on explicit knowledge in declarative memory, which developed through the explanation of rules and conscious attention to input. In contrast, learners with the IE sequence learned English word stress relying more on procedural memory, and they experienced an immersed learning environment and an implicit way of input. Despite potential forgetting, procedural memory often leads to more stable long-term gains than declarative memory (Diekelmann et al., 2009), which may explain the IE group's better persistence and generalization.

Previous studies have suggested a competitive interaction between declarative and procedural memory systems (e.g., Poldrack & Packard, 2003; Ullman, 2020). The "seesaw effect" (Boucher & Anns, 2018; Stocker, 1998) posits that learning in one memory system depressed functionality of the other. In this study, the EI sequence enhanced learners' explicit knowledge via declarative memory, which appeared to inhibit procedural memory's involvement. In contrast, the IE sequence fostered implicit learning and greater reliance on procedural memory for word stress acquisition, thereby hindering the role of declarative memory. This may explain the IE group's better persistence and generalization effects, as procedural memory is known to support consolidation and delayed improvements (Kuriyama et al., 2004). Similar findings in L2 syntax research (Morgan-Short et al., 2012; Faretta-Stutenberg & Morgan-Short, 2018) also highlight the long-term benefits of learning by procedural memory. Our study supported the role of IE instruction sequence but from the angle of L2 word stress identification.

#### **4 Implications, Limitations, and Future Directions**

The practical significance was highlighted for our study. First of all, as a training study participated by high school students, this study provided the pedagogical implications and paved the way for our design to be applied into classroom teaching and learning by teenager learners of an L2. Also, our results and findings are heuristic for the speech treatment of the clinical populations afflicted with speech or hearing disorders. For example, it was found that the hearing-impaired (HI) patients have deficient perception of suprasegmental units of speech (Zhang et al., 2021). The present study offered the normative data regarding the development of young participants' identification of an L2 under blended instruction. The speech-language pathologists could attempt to evaluate whether HI individuals could improve their perceptual performance as normal, typically developing students by providing a more suitable learning method (e.g., the IE sequence). Recent studies also pointed out the degraded speech perception (apart from their production) among individuals who stutter (Zhu et al., 2023b). As the blended instruction facilitates the development of speech perception and, crucially, instruction sequence makes a difference about the perceptual gains, therapists could adopt the blended instruction and determine which instruction sequence would maximize the improvement of stutterers' auditory perception capacities. Being a preliminary endeavor, our study could serve as a baseline for the comparison.

There are some limitations in this study. First, although blended instruction is useful to improve Chinese teenager learners' word stress identification, it remains less known about its effects on word stress production. According to Altmann (2006), despite learners from non-stress languages showing the ceiling perceptual performance, their speech productions were however different. This indicated that speech perception may not necessarily link tightly

to speech production. Second, the absence of intermediate test might restrict our ability to track temporal dynamics of learning (e.g., the learning gains in the explicit or implicit phases and the fluctuations in cognitive load). Future research could incorporate the intermediate test to assess the effects more thoroughly. In addition, setting groups with non-blended instruction (i.e., groups receiving solely implicit or solely explicit instruction) would provide stronger empirical validation (Cai, 2018). Such designs could clarify the distinct contributions of EI and IE sequences to both instructional effects and long-term retention. Third, while this study focused on teenager group, future research should extend to broader age groups to examine instructional sequence effects. For instance, older adults' auditory processing decline (Amieva & Ouvrard, 2020; Bidelman et al., 2019) or children's heightened neuroplasticity (Fandakova & Hartley, 2020; Whelan et al., 2021) may differentially modulate the efficacy of the EI or IE sequences. Future studies could employ fully multimodal perception tests (e.g., Li et al., 2022) and combine identification and discrimination tasks to better explore the perceptual performance of English word stress among learners across the human life span.

## **V Conclusions**

To summarize, our study provides a systematic and comprehensive investigation of the effects of blended instruction on English word stress identification for Chinese teenagers by virtue of GLMM and LMM using R for data analysis. The findings show that blended instruction serves as an effective way to facilitate English word stress identification, and the IE sequence may be more suitable for Chinese teenager learners on account of its better persistence and generalization effects. In addition, the observations in this study support the

prediction of DP model, which proposes that the instruction can affect the reliance on the two memory systems of declarative memory and procedural memory.

## References

- Altmann, H.** (2006). *The Perception and Production of Second Language Stress: A Cross-linguistic Experimental Study*. University of Delaware.
- Altmann, H., & Vogel, I.** (2002). L2 acquisition of stress: The role of L1. In *DGfS Annual Meeting "Multilingualism Today" in Mannheim, Germany* (pp. 69–85).
- Amieva, H., & Ouvrard, C.** (2020). Does treating hearing loss in older adults improve cognitive outcomes? A review. *Journal of Clinical Medicine*, 9(3), 805. <https://doi.org/10.3390/jcm9030805>
- Anttila, A., Dozat, T., Galbraith, D., & Shapiro, N.** (2020). Sentence stress in presidential speeches. *Prosody in Syntactic Encoding, Berlin/Boston: Walter De Gruyter*, 17–50.  
<https://doi.org/10.1515/9783110650532>
- Arthur, D. T., Ullman, M. T., & Earle, F. S.** (2021). Declarative memory predicts phonological processing abilities in adulthood. *Frontiers in Psychology*, 12, 658402. <https://doi.org/10.3389/fpsyg.2021.658402>
- Baayen, R. H., & Milin, P.** (2010). Analyzing reaction times. *International Journal of Psychological Research*, 3(2), 12–28.
- Batel, E.** (2020). Context effect on L2 word recognition: Visual versus auditory modalities. *Journal of Psycholinguistic Research*, 49(2), 223–245. <https://doi.org/10.1007/s10936-019-09683-6>
- Bates, D., Maechler, M., Bolker, B., & Walker, S.** (2020). *Fitting linear mixed-effects models using lme4*. Cornell University. <https://github.com/lme4/lme4/>
- Beckman, M. E.** (1986). *Stress and Non-stress Accent*. Dordrecht, the Netherlands: Foris Publications.

- Bidelman, G. M., Price, C. N., Shen, D., Arnott, S. R., & Alain, C.** (2019). Afferent-efferent connectivity between auditory brainstem and cortex accounts for poorer speech-in-noise comprehension in older adults. *Hearing Research*, 382, 107795. <https://doi.org/10.1016/j.heares.2019.107795>
- Bisson, M. J., Van Heuven, W. J., Conklin, K., & Tunney, R. J.** (2014). Processing of native and foreign language subtitles in films: An eye tracking study. *Applied Psycholinguistics*, 35(2), 399–418. <https://doi.org/10.1017/S0142716412000434>
- Boucher, J., & Anns, S.** (2018). Memory, learning and language in autism spectrum disorder. *Autism & Developmental Language Impairments*, 3, 2396941517742078. <https://doi.org/10.1177/2396941517742078>
- Brill-Schuetz, K., & Morgan-Short, K.** (2014). The role of procedural memory in adult second language acquisition. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).
- Bu, Y., & Zhou, Z.** (2021). A study of Chinese EFL learners' problems with stress Acquisition. *System*, 96, 102392. <https://doi.org/10.1016/j.system.2020.102392>
- Cai, C.** (2018). Effects of Explicit/Implicit Instruction and Frequency on Pronunciation Accuracy of Chinese EFL Learners. *Modern Foreign Languages*, 41(04): 517–528.
- Cai, C.** (2021). Effects of Blended Instruction and Language Proficiency on the Perception Accuracy of English Sentence Stress. *Foreign Languages and Their Teaching*, 1(5): 100–110+150.
- Cebrian, J., Gavaldà, N., Gorba, C., & Carlet, A.** (2024). Differential effects of identification and discrimination training tasks on L2 vowel identification and discrimination. *Studies in Second Language Acquisition*, 46, 1–25. <https://doi.org/10.1017/S0272263124000408>
- Chan, R. K., & Leung, J. H.** (2014). Implicit learning of L2 word stress regularities. *Second Language Research*, 30(4), 463–484. <https://doi.org/10.1177/0267658313510169>

- Chan, R. K., & Leung, J. H.** (2018). Implicit knowledge of lexical stress rules: Evidence from the combined use of subjective and objective awareness measures. *Applied Psycholinguistics*, 39(1), 37–66.  
<https://doi.org/10.1017/S0142716417000376>
- Chen, F., Guo, Q., Deng, Y., Zhu, J., & Zhang, H.** (2023). Development of Mandarin lexical tone identification in noise and its relation with working memory. *Journal of Speech, Language, and Hearing Research*, 66(10), 4100–4116. [https://doi.org/10.1044/2023\\_JSLHR-22-00457](https://doi.org/10.1044/2023_JSLHR-22-00457)
- Chen, H. C.** (2013). Chinese learners' acquisition of English word stress and factors affecting stress assignment. *Linguistics and Education*, 24(4), 545–555. <https://doi.org/10.1016/j.linged.2013.08.003>
- Chen, H. C., Han, Q., & Tian, J. X.** (2025). The Effects of Rule-Based and Acoustic-Perceptual Instruction on Chinese ESL Learners' English Primary Word Stress Placement Accuracy. *International Journal of Applied Linguistics*. In Press. <https://doi.org/10.1111/ijal.12687>
- Chen, H. C., & Tian, J.** (2020). The effects of explicit rule and acoustic-perceptual instructions on Chinese ESL learners' prosodic acquisition of English lexical stress. In *10th International Conference on Speech Prosody 2020*, 833–837. <http://dx.doi.org/10.21437/SpeechProsody.2020-170>
- Chen, Y., Robb, M. P., Gilbert, H. R., & Lerman, J. W.** (2001). A study of sentence stress production in Mandarin speakers of American English. *The Journal of the Acoustical Society of America*, 109(4), 1681–1690. <http://dx.doi.org/10.1121/1.1356023>
- Choi, W., Tong, X., & Samuel, A. G.** (2019). Better than native: Tone language experience enhances English lexical stress discrimination in Cantonese-English bilingual listeners. *Cognition*, 189, 188–192.  
<http://dx.doi.org/10.1016/j.cognition.2019.04.004>
- Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J.** (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research*, 57(4),

1468–1479. [http://dx.doi.org/10.1044/2014\\_JSLHR-L-13-0279](http://dx.doi.org/10.1044/2014_JSLHR-L-13-0279)

**Cope, T. E., Sohoglu, E., Sedley, W., Patterson, K., Jones, P. S., Wiggins, J., ... & Rowe, J. B.** (2017).

Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nature Communications*, 8(1), 2154. <https://doi.org/10.1038/s41467-017-01958-7>

**Couper, G.** (2022). A Conceptual Approach to Teaching L2 Pronunciation: Perception of Word Stress.

*TESL-EJ*, 26(1), n1. <https://doi.org/10.55593/ej.26101a6>

**Davis, S. M., & Kelly, M. H.** (1997). Knowledge of the English noun–verb stress difference by native and nonnative speakers. *Journal of Memory and Language*, 36(3), 445–460.

<https://doi.org/10.1006/jmla.1996.2503>

**Diekelmann, S., Wilhelm, I., & Born, J.** (2009). The whats and whens of sleep-dependent memory

consolidation. *Sleep Medicine Reviews*, 13(5), 309–321. <https://doi.org/10.1016/j.smrv.2008.08.002>

**Durbahn, M., Rodgers, M., & Peters, E.** (2020). The relationship between vocabulary and viewing

comprehension. *System*, 88, 102166. <https://doi.org/10.1016/j.system.2019.102166>

**Eichenbaum, H.** (2004). Hippocampus: cognitive processes and neural representations that underlie

declarative memory. *Neuron*, 44(1), 109–120. <https://doi.org/10.1016/j.neuron.2004.08.028>

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

<https://doi.org/10.1017/S0272263102002024>

**Ellis, N. C.** (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge.

*Studies in Second Language Acquisition*, 27(2), 305–352. <https://doi.org/10.1017/S027226310505014X>

**Ellis, N. C.** (2011). Implicit and explicit SLA and their interface. *Implicit and Explicit Language Learning:*

*Conditions, Processes, and Knowledge in SLA and Bilingualism*, 35, 47.

- Ellis, R.** (2009). Implicit and explicit learning, knowledge and instruction. *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*, 42, 3–25.  
<https://doi.org/10.21832/9781847691767-003>
- Fainman, I., & Tokar, Y.** (2019). Explicit, implicit, and blended vocabulary instruction: Efficiency in an Aviation English Course. *The Collegiate Aviation Review International*, 37(2).  
<https://doi.org/10.22488/okstate.19.100218>
- Fandakova, Y., & Hartley, C. A.** (2020). Mechanisms of learning and plasticity in childhood and adolescence. *Developmental Cognitive Neuroscience*, 42, 100764.  
<https://doi.org/10.1016/j.dcn.2020.100764>
- Faretta-Stutenberg, M., & Morgan-Short, K.** (2018). The interplay of individual differences and context of learning in behavioral and neurocognitive second language development. *Second Language Research*, 34(1), 67–101. <https://doi.org/10.1177/0267658316684903>
- Field, J.** (2005). Intelligibility and the listener: The role of lexical stress. *TESOL quarterly*, 39(3), 399–423.  
<https://doi.org/10.2307/3588487>
- Frost, D.** (2011). Stress and cues to relative prominence in English and French: A perceptual study. *Journal of the International Phonetic Association*, 41(1), 67–84. <https://doi.org/10.1017/S0025100310000253>
- Fukuta, J., & Yamashita, J.** (2023). The complex relationship between conscious/unconscious learning and conscious/unconscious knowledge: The mediating effects of salience in form–meaning connections. *Second Language Research*, 39(2), 425–446. <https://doi.org/10.1177/02676583211044950>
- Gordon, J., Darcy, I., & Ewert, D.** (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*. Aug. 2012. (pp. 194–206). Ames, IA: Iowa State

University.

- Graaff, Rick.** (1997). THE EXPERANTO EXPERIMENT: Effects of Explicit Instruction on Second Language Acquisition. *Studies in Second Language Acquisition*, 19, 249–276.  
<http://dx.doi.org/10.1017/S0272263197002064>
- Grant, K. W., & Seitz, P. F.** (2000). The recognition of isolated words and words in sentences: Individual variability in the use of sentence context. *The Journal of the Acoustical Society of America*, 107(2), 1000–1011. <https://doi.org/10.1121/1.428280>
- Horgues, C.** (2013). French learners of L2 English: Intonation boundaries and the marking of lexical stress. *Research in Language*, 11(1), 41–56. <http://dx.doi.org/10.2478/v10015-012-0006-8>
- Ingram, J.** (2005). Vietnamese acquisition of English word stress. *Tesol Quarterly*, 39(2), 309–319.  
<https://doi.org/10.2307/3588314>
- Ji, X., Zhang, H., Li, A.** (2023). The effect of explicit/implicit training on English intonation acquisition by second language learners. *Foreign Language Teaching and Research*, 55(1), 66–78+159–160.
- Jović, O.** (2016). Durbin-Watson partial least-squares regression applied to MIR data on adulteration with edible oils of different origins. *Food Chemistry*, 213, 791–798.  
<https://doi.org/10.1016/j.foodchem.2016.07.016>
- 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>
- Karatekin, C., Marcus, D. J., & White, T.** (2007). Oculomotor and manual indexes of incidental and intentional spatial sequence learning during middle childhood and adolescence. *Journal of Experimental Child Psychology*, 96(2), 107–130. <https://doi.org/10.1016/j.jecp.2006.05.005>

- Keck, C., Iberri-Shea, G., Tracy-Ventura, N., & Wa-Mbaleka, S.** (2006). Investigating the empirical link between task-based interaction and acquisition. *Synthesizing Research on Language Learning and Teaching*, 13, 91.
- Kim, J. Y.** (2020). Discrepancy between heritage speakers' use of suprasegmental cues in the perception and production of Spanish lexical stress. *Bilingualism: Language and Cognition*, 23(2), 233–250. <https://doi.org/10.1017/S1366728918001220>
- Kim, B., Bishop, J., & Whalen, D. H.** (2025). Lexical Stress in Mandarin Second-Language Speakers of English: An Electromagnetic Articulography Study. *Journal of Speech, Language, and Hearing Research*, 68(3), 1–14. [https://doi.org/10.1044/2024\\_JSLHR-24-00491](https://doi.org/10.1044/2024_JSLHR-24-00491)
- Kuriyama, K., Stickgold, R., & Walker, M. P.** (2004). Sleep-dependent learning and motor-skill complexity. *Learning & Memory*, 11(6), 705–713. <http://www.learnmem.org/cgi/doi/10.1101/lm.76304>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B.** (2017). lmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82(13). <https://doi.org/10.18637/jss.v082.i13>
- Lai, C., Sui, Y., & Yuan, J.** (2010). A corpus study of the prosody of polysyllabic words in Mandarin Chinese. In *Proceedings of Speech Prosody* (Vol. 2010).
- Lee, G., Shin, D. J., & Garcia, M. T. M.** (2019). Perception of lexical stress and sentence focus by Korean-speaking and Spanish-speaking L2 learners of English. *Language Sciences*, 72, 36–49. <https://doi.org/10.1016/j.langsci.2019.01.002>
- Lenth, R. V.** (2016). Least-squares means: the R package lsmeans. *Journal of Statistical Software*, 69, 1–33. <https://doi.org/10.18637/jss.v069.i01>
- Lewis, C., & Deterding, D.** (2018). Word Stress and Pronunciation Teaching in English as a Lingua Franca Contexts. *CATESOL Journal*, 30(1), 161–176.

- Li, M., Chen, X., Zhu, J., & Chen, F.** (2022). Audiovisual mandarin lexical tone perception in quiet and noisy contexts: the influence of visual cues and speech rate. *Journal of Speech, Language, and Hearing Research, 65*(11), 4385–4403. [https://doi.org/10.1044/2022\\_JSLHR-22-00024](https://doi.org/10.1044/2022_JSLHR-22-00024)
- Li, P., Zhang, F., Yu, A., & Zhao, X.** (2020). Language History Questionnaire (LHQ3): An enhanced tool for assessing multilingual experience. *Bilingualism: Language and Cognition, 23*(5), 938–944. <https://doi.org/10.1017/S1366728918001153>
- Lin, C. Y., Wang, M. I. N., Idsardi, W. J., & Xu, Y. I.** (2014). Stress processing in Mandarin and Korean second language learners of English. *Bilingualism: Language and Cognition, 17*(2), 316–346. <https://doi.org/10.1017/S1366728913000333>
- Lindemann, S., Campbell, M. A., Litzenberg, J., & Close Subtirelu, N.** (2016). Explicit and implicit training methods for improving native English speakers' comprehension of nonnative speech. *Journal of Second Language Pronunciation, 2*(1), 93–108. <https://doi.org/10.1075/jslp.2.1.04lin>
- Lukács, Á., Kemény, F., Lum, J. A., & Ullman, M. T.** (2017). Learning and overnight retention in declarative memory in specific language impairment. *PloS One, 12*(1), e0169474. <https://doi.org/10.1371/journal.pone.0169474>
- Ma, Z., Jia, J., Wu, S.** (2022). On the analysis of reaction time data through the use of mixed-effects models: principles and practices. *Foreign Language Learning Theory and Practice, 177*(1), 35–46.
- Macmillan, N. A., & Creelman, C. D.** (2004). *Detection Theory: A User's Guide*. Psychology press.
- Mayer, R. E.** (2014). Incorporating motivation into multimedia learning. *Learning and Instruction, 29*, 171–173. <https://doi.org/10.1016/j.learninstruc.2013.04.003>
- Misyak, J. B., Christiansen, M. H., & Bruce Tomblin, J.** (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science, 2*(1), 138–153.

<https://doi.org/10.1111/j.1756-8765.2009.01072.x>

**Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. T. (2012).** Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, 24(4), 933–947. [https://doi.org/10.1162/jocn\\_a\\_00119](https://doi.org/10.1162/jocn_a_00119)

**Moyer, A. (1999).** Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21(1), 81–108.

<https://doi.org/10.1017/S0272263199001035>

**Nagle, C., & Hiver, P. (2024).** Optimizing second language pronunciation instruction: Replications of Martin and Sippel (2021), Olson and Offerman (2021), and Thomson (2012). *Language Teaching*, 57(3), 419–432. <https://doi.org/10.1017/S0261444823000083>

**Nation, I. S. P., & Beglar, D. (2007).** A vocabulary size test. *The Language Teacher*, 31(7), 9–13.

**Nissen, M. J., & Bullemer, P. (1987).** Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, 19(1), 1–32. [https://doi.org/10.1016/0010-0285\(87\)90002-8](https://doi.org/10.1016/0010-0285(87)90002-8)

**Norris, J. M., & Ortega, L. (2000).** Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528. <https://doi.org/10.1111/0023-8333.00136>

**O'brien, M. G. (2019).** Attending to second language lexical stress: Exploring the roles of metalinguistic awareness and self-assessment. *Language Awareness*, 28(4), 310–328.

<https://doi.org/10.1080/09658416.2019.1625912>

**Ortega-Llebaria, M., & Colantoni, L. (2014).** L2 English intonation: Relations between form-meaning associations, access to meaning, and L1 transfer. *Studies in Second Language Acquisition*, 36(2), 331–353. <https://doi.org/10.1017/S0272263114000011>

**Ortega-Llebaria, M., Gu, H., & Fan, J. (2013).** English speakers' perception of Spanish lexical stress:

Context-driven L2 stress perception. *Journal of Phonetics*, 41(3-4), 186–197.

<https://doi.org/10.1016/j.wocn.2013.01.006>

**Ortín, R., & Simonet, M.** (2023). Perceptual sensitivity to stress in native English speakers learning Spanish as a second language. *Laboratory Phonology*, 14, 1–41. <https://doi.org/10.16995/labphon.7978>

**Parlak, Ö., Ziegler, N., & Rabea, R.** (2025). Implicit feedback and lexical stress: Development of perception and production in a classroom environment. *Journal of Second Language Pronunciation*, (Online Published). <https://doi.org/10.1075/jslp.24050.par>.

**Peters, E., & Muñoz, C.** (2020). Introduction to special issue Language learning from multimodal input. *Studies in Second Language Acquisition*, 42(3), 489–497.

**Poldrack, R. A., & Packard, M. G.** (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia*, 41(3), 245–251.

[https://doi.org/10.1016/S0028-3932\(02\)00157-4](https://doi.org/10.1016/S0028-3932(02)00157-4)

**Pujadas, G., & Muñoz, C.** (2020). Examining adolescent EFL learners' TV viewing comprehension through captions and subtitles. *Studies in Second Language Acquisition*, 42(3), 551–575.

<https://doi.org/10.1017/S0272263120000042>

**Qin, Z., Chien, Y. F., & Tremblay, A.** (2017). Processing of word-level stress by Mandarin-speaking second language learners of English. *Applied Psycholinguistics*, 38(3), 541–570.

<https://doi.org/10.1017/S0142716416000321>

**R Core Team.** (2023). “R: A language and environment for statistical computing,” R Foundation for Statistical Computing. <http://www.R-project.org/>

**Reifegerste, J., Veríssimo, J., Rugg, M. D., Pullman, M. Y., Babcock, L., Gleib, D. A., ... & Ullman, M. T.** (2021). Early-life education may help bolster declarative memory in old age, especially for women.

*Aging, Neuropsychology, and Cognition*, 28(2), 218–252.

<https://doi.org/10.1080/13825585.2020.1736497>

**Ripley, B., Hornik, K., Gebhardt, A., & Firth, D.** (2012). Functions and datasets to support Venables and

Ripley, *Modern Applied Statistics with S* (2002): package “MASS.” Accessed through

<https://cran.r-project.org/>

**Robinson, P.** (1997). Individual differences and the fundamental similarity of implicit and explicit adult

second language learning. *Language Learning*, 47(1), 45–99.

<https://doi.org/10.1111/0023-8333.21997002>

**Rodgers, M. P., & Webb, S.** (2017). The effects of captions on EFL learners' comprehension of

English-language television programs. *Calico Journal*, 34(1), 20–38. <https://doi.org/10.1558/cj.29522>

**Roehr-Brackin, K.** (2014). Explicit knowledge and processes from a usage-based perspective: The

developmental trajectory of an instructed L2 learner. *Language Learning*, 64(4), 771–808.

<https://doi.org/10.1111/lang.12081>

**Ruchkin, D. S., Grafman, J., Cameron, K., & Berndt, R. S.** (2003). Working memory retention systems:

A state of activated long-term memory. *Behavioral and Brain Sciences*, 26(6), 709–728.

<https://doi.org/10.1017/S0140525X03000165>

**Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L.** (1999). Statistical learning of tone

sequences by human infants and adults. *Cognition*, 70(1), 27–52.

[https://doi.org/10.1016/S0010-0277\(98\)00075-4](https://doi.org/10.1016/S0010-0277(98)00075-4)

**Saito, K.** (2019). The role of aptitude in second language segmental learning: The case of Japanese learners'

English/r/pronunciation attainment in classroom settings. *Applied Psycholinguistics*, 40(1), 183–204.

<https://doi.org/10.1017/S0142716418000528>

- 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, Y., & Saito, K.** (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research*, 21(5), 589–608. <https://doi.org/10.1177/1362168816643111>
- Schwab, S., & Dellwo, V.** (2022). Explicit versus non-explicit prosodic training in the learning of Spanish L2 stress contrasts by French listeners. *Journal of Second Language Studies*, 5(2), 266–306. <https://doi.org/10.1075/jsls.21017.sch>
- Shanks, D. R.** (2003). Attention and awareness in “implicit” sequence learning. In *Attention and implicit learning* (pp. 11–42). John Benjamins Publishing Company.
- Silveira, R.** (2002). Pronunciation instruction classroom practice and empirical research. *Revista Linguagem & Ensino*, 5(1), 93–126. <https://doi.org/10.15210/rle.v5i1.15552>
- Sippel, L., & Martin, I. A.** (2023). Immediate and long-term improvement in lexical stress perception: the role of teacher and peer feedback. *International Review of Applied Linguistics in Language Teaching*, 61(3), 1173–1195.
- Spada, N., & Tomita, Y.** (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308. <https://doi.org/10.1111/j.1467-9922.2010.00562.x>
- Stocker, T. F.** (1998). The seesaw effect. *Science*, 282(5386), 61–62. <https://doi.org/10.1126/science.282.5386.61>
- Sweller, J., van Merriënboer, J. J., & Paas, F.** (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261–292.

- Syndicate, U.** (2001). *Quick Placement Test*. Oxford: Oxford University Press.
- Uchihara, T., Webb, S., Saito, K., & Trofimovich, P.** (2022). The effects of talker variability and frequency of exposure on the acquisition of spoken word knowledge. *Studies in Second Language Acquisition*, 44(2), 357–380. <https://doi.org/10.1017/S0272263121000218>
- Ullman, M. T.** (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, 2(10), 717–726. <http://dx.doi.org/10.1038/35094573>
- Ullman, M. T.** (2016). The Declarative/Procedural Model: A Neurobiological Model of Language Learning, Knowledge, and Use. In *Neurobiology of Language* (pp. 953–968). Academic Press.
- Ullman, M. T.** (2020). The declarative/procedural model: A neurobiologically motivated theory of first and second language 1. In *Theories in Second Language Acquisition* (pp. 128–161). Routledge.
- Ullman, M. T., & Lovelett, J. T.** (2018). Implications of the declarative/procedural model for improving second language learning: The role of memory enhancement techniques. *Second Language Research*, 34(1), 39–65. <https://doi.org/10.1177/0267658316675195>
- VanPatten, B., & Cadierno, T.** (1993). Explicit instruction and input processing. *Studies in Second Language Acquisition*, 15(2), 225–243. <http://dx.doi.org/10.1017/S0272263100011979>
- Wang, Q.** (2008). *Perception of English Stress by Mandarin Chinese Learners of English: An Acoustic Study*. University of Victoria.
- Wayland, R., Landfair, D., Li, B., & Guion, S. G.** (2006). Native Thai speakers' acquisition of English word stress patterns. *Journal of Psycholinguistic Research*, 35, 285–304. <https://doi.org/10.1007/s10936-006-9016-9>
- Whelan, B. M., Theodoros, D., McMahon, K. L., Copland, D., Aldridge, D., & Campbell, J.** (2021). Substrates of speech treatment-induced neuroplasticity in adults and children with motor speech

- disorders: A systematic scoping review of neuroimaging evidence. *International Journal of Speech-Language Pathology*, 23(6), 579–592. <https://doi.org/10.1080/17549507.2021.1908425>
- Xiong, Ziyu.** (2019). *The new function of xPerception* [OL].  
<https://mp.weixin.qq.com/s/apjADRgQj1SREycqPnA5Kw> (accessed 11/12/2019).
- Yenkimaleki, M., & Van Heuven, V. J.** (2016). Explicit teaching of segmentals versus suprasegmentals: which would yield better listening comprehension skills for interpreter trainees? An experimental study. *British Journal of English Linguistics*, 4(6), 11–22.
- Yu, V. Y., & Andruski, J. E.** (2010). A cross-language study of perception of lexical stress in English. *Journal of Psycholinguistic Research*, 39, 323–344. <https://doi.org/10.1007/s10936-009-9142-2>
- Zibin, A., & Altaxhaineh, A. R. M.** (2019). The effect of blended learning on the development of clause combining as an aspect of the acquisition of written discourse by Jordanian learners of English as a foreign language. *Journal of Computer Assisted Learning*, 35(2), 256–267.  
<https://doi.org/10.1111/jcal.12327>
- Zhang, H., Ding, H., & Zhang, Y.** (2021). High variability phonetic training benefits lexical tone perception: An investigation on Mandarin-speaking pediatric cochlear implant users. *Journal of Speech, Language, and Hearing Research*, 64(6), 2070–2084. [https://doi.org/10.1044/2021\\_JSLHR-20-00631](https://doi.org/10.1044/2021_JSLHR-20-00631)
- Zhang, L., Zhu, J., Shao, J., & Zhang, C.** (2024). The combined effects of bilingualism and musicianship on listeners' perception of non-native lexical tones. *Speech Communication*, 165, 103147.  
<https://doi.org/10.1016/j.specom.2024.103147>
- Zhang, Y., Nissen, S. L., & Francis, A. L.** (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America*, 123(6), 4498–4513. <http://dx.doi.org/10.1121/1.2902165>

**Zhu, J., Chen, X., Chen, F., Zhang, C., Shao, J., & Wiener, S. (2023a).** Tone deafness in music does not preclude distributional learning of nonnative tonal languages in individuals with congenital amusia.

*Journal of Speech, Language, and Hearing Research*, 66(7), 2461–2477.

[https://doi.org/10.1044/2023\\_JSLHR-22-00572](https://doi.org/10.1044/2023_JSLHR-22-00572)

**Zhu, J., Shao, J., Chen, F., Zhang, C., & Wiener, S. (2023b).** Statistical information affects spoken word recognition of tone languages in stutterers: Evidence from an auditory-perceptual gating study. *Journal of Speech, Language, and Hearing Research*, 66(9), 3382–3398.

[https://doi.org/10.1044/2023\\_JSLHR-23-00123](https://doi.org/10.1044/2023_JSLHR-23-00123)

**Zubizarreta, M. L., & Nava, E. (2011).** Encoding discourse-based meaning: Prosody vs. syntax.

Implications for second language acquisition. *Lingua*, 121(4), 652–669.

<https://doi.org/10.1016/j.lingua.2010.06.013>

## Tables

Table 1. Demographic characteristics of EI and IE groups.

Subject Information	EI Group	IE Group
No. of Participants (Female)	40 (20)	40 (20)
Age [in years; (SD)]	15.1 (0.71)	15.2 (0.60)
Years of Learning English (SD)	6.04 (0.26)	6.12 (0.39)
Quick Placement Test (SD)	30.8 (5.22)	30.7 (5.65)
Vocabulary Size Test (SD)	92.2 (6.88)	92.4 (7.25)
Declarative Memory [ $d'$ score; (SD)]	0.45 (0.19)	0.655 (0.21)
Procedural Memory [correct RT; (SD)]	463.4 (131)	605.4 (177)

*Note.* EI = explicit and then implicit instruction; IE = implicit and then explicit instruction; RT = reaction time.

Table 2. Acoustic parameters between the first syllable and the second syllable with an example of “conduct.”

Word	Parameters	The first syllable	The second syllable
CONduct	F0 (Hz)	125.64	99.25
	Intensity (dB)	74.54	68.77
	Vowel Duration (ms)	152.47	131.18
	F2 (Hz)	1734.66	1361.47
conDUCT	F0 (Hz)	114.32	126.93
	Intensity (dB)	62.87	75.41
	Vowel Duration (ms)	103.84	116.50
	F2 (Hz)	1585.45	1712.77

*Note.* F0 = fundamental frequency; F2 = the second formant.

## Figures

Figure 1. The experimental procedure designed in this study. EI = explicit and then implicit instruction; IE = implicit and then explicit instruction.

Figure 2. Mean accuracy of English word stress perception across pretest, immediate post-test, and delayed post-test on trained and untrained words by EI and IE groups. EI = explicit and then implicit instruction; IE = implicit and then explicit instruction.

Figure 3. Mean RT of English word stress perception across pretest, immediate post-test, and delayed post-test on trained and untrained words by EI and IE groups. Error bars represent 95% confidence intervals. EI = explicit and then implicit instruction; IE = implicit and then explicit instruction.

Figure 4. The relations between accuracy and declarative memory in EI and IE groups. The shadow area represents the 95% confidence interval of the best-fit line from regression models. EI = explicit and then implicit instruction; IE = implicit and then explicit instruction.

Figure 5. The relations between accuracy and procedural memory in EI and IE groups. The shadow area represents the 95% confidence interval of the best-fit line from regression models. EI = explicit and then implicit instruction; IE = implicit and then explicit instruction.