8-21-2024 10:00 AM

Learning the pronunciation of English words from textual input: Should we listen first?

Liwen Situ, *Western University*

Supervisor: Boers, Frank, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Master of Arts degree in Education

Abstract

This study investigated factors influencing incidental English word pronunciation acquisition by upper-intermediate L2 learners through exposure to spoken discourse. Due to inconsistent English spelling-sound correspondences, silent reading is likely to leave learners with inaccurate pronunciations. This study explored whether these inaccuracies could be easily corrected through listening. Two sequences were compared: silent reading followed by listening and listening followed by silent reading.

In a counterbalanced within-participant design, 50 upper-intermediate ESL learners at a research-intensive University in Ontario engaged with a text containing 16 target words. The text was divided into to parts. Participants either read a part silently, then aloud, followed by listening, or they listened first, then read silently and aloud. The sequence was reversed for the other part of the text. Post-tests assessed pronunciation improvements and interviews explored individual differences.

The results indicated that a single audio exposure was insufficient for accurate pronunciation acquisition. Both the trial-and-error and retrieval approaches yielded comparable final outcomes. However, the Input-Output-Input sequence (listening, reading, and listening again) showed potential as a more effective teaching strategy, combining the benefits of both approaches to enhance learning outcomes.

Keywords

pronunciation; trial-and-error; retrieval; noticing; output hypothesis; memory; learning styles

Summary for Lay Audience

English pronunciation is notoriously unpredictable. For example, words like "*through*," "*thought*," "*though*," "*tough*," and "*plough*" all share the *-ough* spelling but are pronounced differently. Similarly, stress placement varies: in "*precursor*" and "*pretentious*," stress is on the second syllable, while in "*predator*" and "*prevalent*," it's on the first. These inconsistencies can lead to incorrect representations of the words' pronunciation in learners' minds during silent reading. This study explores whether these inaccuracies can be corrected when learners hear the words spoken later.

This issue is important because many English learners primarily encounter the language through reading. If incorrect pronunciations from reading are not easily corrected by listening, it might be better to expose learners to spoken language first. On the other hand, silent reading might spark learners' curiosity about pronunciations, thus making learners attend to the spoken word forms when they hear them at a later point in time.

The study involved high-intermediate ESL learners at a research-intensive University in Ontario, using a design where each participant experienced both learning sequences. Participants engaged with a text containing 16 commonly mispronounced words. They either read one half of the text silently, then aloud, followed by listening, or listened first, then read silently and aloud. For the other half of the text, the sequence was reversed. This design allowed comparison of pronunciation accuracy before and after listening. Participants also completed a learning style survey and were interviewed about their experiences and whether they imagined the spoken form of the words during silent reading.

The results showed that listening to a word just once was not rarely enough for learners to pronounce it correctly. Both methods - guessing the pronunciation and then hearing the correct version, or hearing it first and then reading it - led to similar results. A more effective approach may be a combination of both: listening to the word, reading it, and then listening again.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Frank Boers. Conducting this study has been a challenging journey, especially as a first timer. Your continuous encouragement was like a lighthouse in the darkness. Your comments and suggestions always pushed me to think deeper. You have been a role model in work ethics and commitment to academic excellence, I truly appreciate your kind understanding, valuable insights and tireless patience.

I am also grateful to my peers in the Vocabulary Group at Western University, we have the best team! Thanks for all the selfless advice, support and hugs! Special thanks to Alyssa and Emi, who patiently guided me through data analysis and shared resources. And I will never forget the days and nights we fought for our thesis at the graduate lounge, Lihan and Tina, you are the best stay-up-late-mates.

A heartfelt thank you to the 50 participants who dedicated their time and shared their beliefs, thoughts, and experiences. Your cooperation was invaluable to this research.

I would like to thank the professors and staff in the Faculty of Education at Western University for their assistance with recruitment and data collection. Western University has provided excellent support through its lectures, mentors, and facilities, including 24-hour access to the faculty building.

To my parents, my first teachers, thank you for always emphasizing the importance of education and for your endless love. Your financial support allowed me to study overseas, which has been the most precious gift in my life.

Special thanks to my teachers: Huiping Huang (黄惠萍), who instilled in me a love for reading and a competitive spirit; Song Dai (戴松), who taught me to be a lifelong learner and encouraged me to create my own story, dream, and set goals; and Kelly Zhang (章莉老师), whose passion for education inspired me to pursue it as my lifelong career.

Nikki, thank you for helping me find my true calling.

Jeffrey, whom I met at the beginning of my master's journey, and now my soon-to-be husband (our wedding is 5 days away when I am writing the acknowledgement), thank you for your unwavering love and support as I pursued my dream.

Lastly, to my family and friends in Canada and China, no matter how far and how long we are apart, I know our hearts are together. Thanks for all your love!

Table of Contents

List of Tables

List of Appendices

Chapter 1 Introduction

"To guess or not to guess, this is the question." Through my learning and teaching experience, I have witnessed that the two opposite teaching sequences, trial-and-error followed by feedback versus studying followed by retrieval practice, have both been very popular in many subjects in school. The trial-and-error sequence encourages students to infer or guess new knowledge themselves before the teacher confirms, refines, or rectifies their inferences or guesses. By contrast, in the retrieval sequence students are given the correct information from the start and they are then given the task to retrieve this knowledge from memory. The trial-and-error method is believed to be beneficial for learning because it piques students' curiosity about the correct information. On the downside, there may be a risk that wrong guesses linger in the students' memory and interfere with later recall of the correct information. Presenting the correct information from the start minimizes this risk. Given that both approaches are common, I became curious to investigate which one was more effective. Supportive evidence was found in the existing literature for both approaches regarding diverse target learning contents in the domain of cognitive psychology, as well as language learning (e.g., Haight et al., 2007; Herron & Tomasello, 1992; Vogel et al., 2011; Wang et al., 2022). However, few studies have focused on pronunciation.

Pronunciation is one of the most important aspects of vocabulary knowledge (Nation, 2013), and sufficiently accurate and intelligible oral production of words is crucial for L2 learners to achieve successful spoken communication (Uchihara, 2022). Instructional approaches to pronunciation are necessary because it is commonly found that L2 learners possess weaker knowledge of the spoken forms of English words than their written forms (e.g., Milton & Hopkins, 2006; Uchihara & Harada, 2018). One of the explanations for this phenomenon is that English is not an *alphabetic ideal* writing system with a consistent one-to-one correspondence between phonemes and graphemes, in comparison to languages like Spanish and Czech. There are inconsistent spellings (many-to-one letter-sound and sound-letter correspondences), irregular spelling patterns and silent letters in the English language (Caravolas, 2004). Zwier and Boers (2023) point out, for instance, that the vowel sounds in *meet, meat, seize, siege,* and *pristine* are

the same, although the spellings differ. On the other hand, *through, thought, though, tough,* and *plough* share the same *-ough* spelling, but each is pronounced differently. The letter *k* is not pronounced in the words *know, knight*, and *knee*, nor do the words *bomb, doubt,* and *debt* include the letter *b* sound in their pronunciation. In a similar vein, it can be hard to predict which syllable of a bi-syllabic or multi-syllabic word receives primary stress in English. In *precursor* and *pretentious,* the primary stress is on the second syllable, whereas in *predator* and *prevalent,* the stress is on the first syllable. The first syllable is stressed in *necessary*, but the second syllable is stressed in *necessity* and the third syllable is stressed in *necessarily.* This degree of variability is different from languages such as French, where the primary word stress in multi-syllabic words is systematically on the final syllable, and Italian, where it is on the penultimate syllable. In a nutshell, there are no straightforward rules that help learners make reliable predictions of how certain words in English are pronounced. Rather than being a matter of "system" or "rule" learning, where learners apply knowledge that a given grapheme systematically represents a certain phoneme and where they apply knowledge of a systematic prosodic pattern, learning the pronunciation of many English words needs to happen at the level of "item" learning, where the phonological form of individual words is processed, stored, and retrieved holistically, that is, through establishing larger-unit sound-spelling correspondences (Kessler, 2009). This issue is potentially problematic, because in many contexts written texts remain the primary source for L2 English learning. If a learner encounters new vocabulary through silent reading and creates a mental representation of the spoken form of a word, then this imagined pronunciation may not correspond to the actual pronunciation of the word due to the irregular correspondence between written and spoken forms in English.

While this dissertation is about L2 learning, it is worth mentioning that these irregularities present challenges not only for L2 learners but also for native English speakers (L1). L1 speakers may first encounter certain low-frequency words in written form before hearing them, leading to potential mispronunciations. However, due to their extensive exposure to spoken language, L1 users are generally better equipped to make accurate predictions about how words are pronounced, because they can draw on implicit analogies with a large reservoir of previously learned spoken word forms.

Similarly, highly proficient L2 learners may benefit from a greater familiarity with spoken word forms, allowing them to make more accurate predictions than less proficient learners. While both groups can use prior knowledge to navigate the irregularities of English pronunciation, they are not immune to errors. This makes the study relevant, as it investigates whether listening before reading can help prevent incorrect imagined pronunciations.

If a learner imagines the spoken form of new words encountered during silent reading, and then receives implicit feedback if they later hear the correct pronunciation of the words in aural input, this can be considered an example of learning through trial and error. In this scenario, reading a text first may make learners notice that they do not know how to pronounce certain words yet. This may happen especially if they read the text aloud. Having noticed a gap in their knowledge, they may then pay attention to these words when they hear them in speech. Alternatively, learners may already have heard the new words in aural input before they get to meet them in written texts. If so, the written text may prompt reactivation in their minds of the spoken form. This scenario is expected to be less error-prone and could be considered a case of retrieval from memory of previously acquired knowledge. The objective of this proposed study is to compare the two scenarios just described, and thus to investigate whether the trial-and-error sequence or the retrieval sequence is more effective for the learning of easily mispronounced English words. It is worth mentioning that studies of the relative benefits of trial-and-error versus retrieval procedures have typically been conducted so far in contexts of intentional learning, where participants are given explicit instructions to either use inferencing/guessing strategies or to study items and then recall them. The present study, however, is situated in the realm of incidental acquisition, where the learners are more likely to engage primarily with the content of discourse (i.e., what is said) rather than the precise wording used to express this content (i.e., how it is said).

The participants in this experimental study were high-intermediate ESL learners, who received the same written and aural input texts but in a different sequence, after which they took a test on the pronunciation of pre-selected target words (see Methodology for details). They were also interviewed about their learning style and

language learning habits. The findings from this investigation may inform learners, practitioners, and researchers about potential ways of improving L2 English pronunciation learning.

Chapter 2 Literature Review

2.1 Inductive Versus Deductive Learning

The distinction between inductive and deductive learning has been applied mostly in the domain of L2 grammar learning. The inductive teaching approach follows the sequence that firstly presents students with examples and encourages them to discover the recurring pattern or rule, after which the pattern or rule is explained to them, so the students can confirm, finetune, or rectify their hypotheses (Boers, 2021). Some forms of the inductive approach require students to infer the rule without assistance (Rosa & O'Neill, 1999; Shaffer, 1989), some provide guiding questions (Herron & Tomasello, 1992); some use made-up and decontextualized examples of the target pattern, while others prefer examples drawn from authentic discourse (Adair-Hauck et al., 2005). The deductive approach, by contrast, presents the pattern or rule to students directly and explicitly, and then offers opportunities to students to apply it through practice (Boers, 2021).

Researchers have been interested in comparing the effectiveness and efficiency of the two approaches mostly with a focus on grammar, and supportive evidence has been reported for both. A small-scale quasi-experimental study conducted by Herron and Tomasello (1992) in an L2 French course on learning a series of French grammar patterns suggested that the inductive method led to better learning outcomes in comparison to the deductive method. However, this may be due to how the deductive approach was designed. The participants in this condition practiced the grammar rules through shallow repetition drill exercises that did not require them to retrieve the newly learned rules from memory. Cerezo et al.'s (2016) mixed-method study on L2 learning of the complex Spanish *gustar* grammar pattern also showed superior retention of the target grammar under the inductive condition. It needs to be said, however, that the inductive approach was implemented in this study in the form of an interactive video game, while the deductive approach was implemented in a teacher-fronted traditional classroom. There was also a difference between the two learning conditions in the amount of time invested. On the other hand, in an experimental study by Robinson (1996) on learning two English grammatical structures (one regarded easier than the other), the learning outcomes for the deductive

method were better. More recent studies, Haight et al. (2007) and its replication study, Vogel et al. (2011), have found only partial support for the inductive approach. Both studies taught half of the target grammar patterns in a French course deductively, and the other half inductively. The immediate post-tests at the end of each lesson indicated an advantage for the inductive method, but an end-of-term test showed no significant difference between the two methods. Interestingly, an exit questionnaire revealed that the students predominantly preferred the deductive method.

As mentioned previously, learning the pronunciation of English words is to a considerable extent a matter of item learning rather than rule learning (or system learning). The literature that discusses the effectiveness of inductive and deductive approaches focuses mostly on learners' ability to discover rules and whether this discovery learning leads to better long-term outcomes than a "rules-given" approach. However, rule (or system) learning is likely to involve different kinds of cognitive processes than item learning. These studies remain very important in understanding the nature and effectiveness of discovery learning, but the notions of trial-and-error and retrieval practice may be more appropriate theoretical frameworks to interpret the outcomes of interventions to help learners to remember the spoken form of English words.

2.2 Trial-And-Error Versus Retrieval Practice

The trial-and-error approach is a similar concept to the inductive method in the sense that learners proffer hypotheses before receiving feedback. Errors are welcome and it is believed that learning occurs through recognizing and correcting one's errors. The main difference between the trial-and-error approach and the inductive method as outlined above is that trial and error may concern single items (including the pronunciation of a particular word), while inductive learning typically concerns discovering a rule by analyzing a considerable number of examples which exhibit the rule. The conclusion reached through inductive reasoning then applies to additional, new cases where the same rule is applied. A similar process of discovering a general pattern based on plenty of examples may apply to pronunciation as well, provided there is sufficient systematicity. In the case of French, for example, one may learn through inductive processes that the written letter *h-* (as in *hazard*) is not pronounced, that the written plural *-s* (as in *pommes*)

is not pronounced, and that word stress is on the final syllable (as in *restaurant*). As explained above, this inductive "rule learning" is far less straightforward when it comes to learning the spoken form of English words because there is far less systematicity. For example, learning the recurring vowel in *touch* and *tough* does not transfer to the pronunciation of *couch*.

Retrieval practice after a study episode resembles the deductive method, and is less error prone, simply because the learners are provided with the "correct" information from the start. Again, the term retrieval practice rather than deductive learning will be adopted in the present study because my research interest lies with the pronunciation of English lexical items rather than "rule" learning.

Akin to findings in examining whether the inductive or the deductive is more effective, there is empirical evidence to support both the trial-and-error and the retrieval-practice approach. Researchers have claimed that the trial-and-error approach leads to better learning outcomes for the following reasons:

1) According to the notion of *desirable difficulties* (Bjork & Bjork, 2011) there is a long-term advantage for learners if the learning task is challenging (Karpicke et al., 2014; Roediger & Butler, 2011; Smith et al., 2013). Generating an answer or solving a problem is more challenging than being spoon-fed the information. This is also known as the *generation effect* (Metcalfe, 2017).

2) Learners' involvement in discovery learning helps to foster learning strategies, which may in turn foster learner autonomy (e.g., Grimaldi & Karpicke, 2012; Huelser & Metcalfe, 2012; Kornell et al., 2009).

However, there is also mounting evidence supporting the approach where learners are directly presented with the material to be studied and are subsequently asked to retrieve the new knowledge from memory. Two experiments by Warmington et al. (2013) and Warmington and Hitch (2014) compared the trial-and-error approach and the retrieval practice approach, with learning the meanings of new individual words (presented aurally) as the research focus. The results of the posttests showed better word recall for the

retrieval practice condition. Another direct comparison of the two approaches was an experimental study by Strong and Boers (2019), about the meanings of phrasal verbs. The posttests indicated significantly better results for the retrieval practice condition.

The effectiveness of learning through trial and error may depend on whether the learners' guesses or inferences are accurate. A study by Wang et al. (2021) in which EFL learners were asked to guess the meaning of idioms before the meanings were given to them revealed that wrong guesses can linger in the memory. In another study regarding idioms, Yu and Boers (2023) found that meaning inferencing led to better recall of the meanings than simply giving the meanings to learners only when the inferences were correct from the start. A study about exercises on verb-noun collocations (Li, 2023) also found that it was when learners chose the right response from the start (i.e., trial without error) that the learning gains were relatively good.

A possible interpretation of the conflicting findings in the existing literature is that the kind of learning contents pose different challenges. The following diverse items have been explored in the literature: English words, obsolete English words or words from a language that the participants had no prior knowledge of, phrasal verbs, idioms, grammatical structures in French and Spanish, and general subject knowledge. To the best of my knowledge, there have been no studies about pronunciation learning from exposure to L2 discourse that have directly compared trial-and-error procedures to retrieval procedures. This study aims to fill this gap, and the findings may inform EFL/ESL learners and teachers about whether they should give precedence to trial-and-error or retrieval practice to foster knowledge of the pronunciation of new words.

The trial-and-error procedure is operationalized here by having students first read a text including new words silently, then aloud, and having them next listen to an audio-recording of the text (which will serve as implicit feedback). The retrieval condition is operationalized by having the students first listen to the audio-recording before they read the texts first in silence and then aloud. After these two activity sequences, the students were given a post-test where they were asked to read aloud 16 preselected words from the text.

2.3 Output Hypothesis & Involvement Load Hypothesis

The hypothesis that the trial-and-error learning sequence may induce better English word pronunciation acquisition is supported by Swain (1995)'s *Output Hypothesis*. When learners follow trial-and-error, they are pushed to notice what is missing from their own L2 repertoire to pronounce the words. These knowledge gaps may prompt learners' attention to the new words' pronunciation when they next have the chance to listen to them in speech.

Another theory that advocates the advantage of the trial-and-error learning sequence in English word pronunciation uptake is *Involvement Load Hypothesis* (Hulstijn & Laufer, 2001). This hypothesis suggests that a learning scenario including three components of engagement, notably need, search and evaluation, is beneficial for word acquisition from text. In the trial-and-error condition of this study, learners may experience a need to acquire the pronunciation of the new words during reading. These deficiencies will urge them to search for the correct pronunciation when they listen to the audio next. Hearing the words gives them the opportunity to evaluate their first-attempted pronunciation by comparing it with what they capture in the audio, potentially leading to a more accurate pronunciation.

2.4 Pronunciation Research

L2 pronunciation has been considered a challenging issue for learners and teachers for centuries, particularly for adult learners. However, the amount of empirical pronunciation research was limited before 2005 (Thomson & Derwing, 2015). Recent decades have witnessed a notable surge in researchers' interests in L2 pronunciation acquisition (Levis & Sonsaat, 2020). The focus has often been on how to teach learners to articulate specific phonetic features (e.g. Derwing et al., 1998; Saalfeld, 2011; Saito, 2012), which typically requires extensive practice before learners can achieve successful oral production for communication purposes.

This study, however, examines another under-researched factor that may also impede pronunciation performance. As mentioned in the introduction, English words have

inherently inconsistent spelling-sound correspondences, creating an additional memory burden for learners who might struggle to infer the accurate pronunciation from the spelling alone. Failure to choose the correct vowels, consonants, or stress placement can affect intelligibility. However, these difficulties can often be overcome through noticing and correction, which is typically easier than learning the articulation of new phonemes.

In this study, I assume that learners have already mastered the articulation of the target sounds and can apply them in extemporaneous speech production. If their pronunciation is erroneous, it is due to incorrect choices among the potential options. Once learners hear the correct pronunciation of a target word, they should be able to adjust their pronunciation accordingly.

Regarding pronunciation rating, more and more researchers now argue that the primary goal of pronunciation research and instruction should be to enhance learners' understandability (e.g. Derwing & Munro, 2005; Jenkins, 2000; Levis, 2005; Thomson & Derwing, 2015). In line with this perspective, we focus on assessing the intelligibility of learners' word pronunciation (rather than the attainment of a "native-like" accent), which is defined as the degree to which a listener can comprehend the words that the speaker intends to communicate (Derwing et al., 2022).

2.5 Learning Styles

Learning styles are a widely acknowledged and accepted concept for L2 teaching practitioners and researchers (Saiz Aja, 2022). Among numerous definitions stated by scholars over the years, Reid's (1995) definition has been widely used (e.g., Dörnyei, 2005; Dörnyei & Ryan, 2015; Griffiths, 2012). According to this definition, individual learners have their preferred way to absorb, process, and retain new knowledge. Various classifications of learning styles have been proposed, but scholars agree on at least three categories - visual, auditory, and tactile/kinesthetic. Learners whose learning style is visual prefer input from sight, such as pictures, charts, written information, objects, etc. (Dörnyei, 2005). Auditory learners prefer receiving information by listening. The tactile/kinesthetic learning style describes learners who learn the best through physical movement and touch, for instance, hands-on activities, role-play, experiments, and so on

(Saiz Aja, 2022). It is plausible that auditory learners are more likely than other learners to imagine the spoken form of words during silent reading. If so, they are more prone to creating mental representations of such spoken forms that will later need to be rectified. Put differently, students with this learning style may be better served by presenting them with aural input first, especially because they are likely to take in the spoken form of the words. By comparison, learners who are less likely to create mental representations of spoken word forms during silent reading may experience less interference from such representations as they learn the actual pronunciation of these words. Because the effects of trial-and-error and retrieval procedures examined in this study may interact with individuals' learning styles, the study also collected information about the participants' learning styles.

Chapter 3 Research Questions

The research questions addressed in this study are as follows:

a) Can learners acquire word pronunciations from one incidental exposure to an audio-recording?

b) Is the trial-and-error or retrieval procedure the more effective approach for learning the pronunciation of new English words?

c) Do individual differences, including differences in learning styles, affect the learning outcomes, according to the participants' retrospective verbal reports?

Chapter 4 Methodology

4.1 Participants

Fifty L2 students enrolled at a research-intensive University in Ontario participated in this study and completed all required tasks. The majority were enrolled in a course-based graduate program in TESOL (n = 33), the others were mostly graduate students in social sciences and arts and humanities (n = 13), except two graduate students who were majoring in medical science and computer science, respectively. Two senior-year undergraduate students in social sciences also participated in the study. The participants' mean length of time living in an English-speaking environment was two years. More than seventy-five percent (n = 38) of them had spent over six months in Canada by the time of this study, while sixty percent (n = 30) of them had arrived in Canada less than one year ago. As regards their L1s, Mandarin Chinese was predominant. Other first languages spoken included Portuguese (n = 1), Farsi (n = 1), Urdu (n = 1), Malayalam (n = 2), and Kannada (n = 1). The latter two languages are spoken in south Indian states.

To gain admission to the university programs, students had to achieve a minimum overall score of 6.5, with no single component (speaking, listening, reading, or writing) lower than 6.0, on the IELTS exam or equivalent tests. Such scores indicate that our participants were at least upper-intermediate L2 learners based on the *Common European Framework of Reference* (CEFR), although their level of proficiency inevitably varied. Therefore, a within-participant design was adopted to prevent any potential confounding factors stemming from individual differences regarding English proficiency.

This study received approval from the Research Ethics Board of the university. Each participant attended one data collection session and received a $25 gift card as compensation with the assurance that their participation was voluntary and would not affect their course grades in any way. All participants provided written consent for their data to be used for research purposes.

4.2 Intervention Material

4.2.1 Article

A modified popular-scientific text and an audio-recording of this text were used as the intervention material. The selection of the original article was based on three criteria: length, reader-interest, and the number of target words. First, since the participants were required to read/listen a total of four times over the course of one meeting, it was crucial to choose an article that is of a proper length and about an interesting topic to maintain participants' attention. Second, to be able to examine if new pronunciation learning occurred, the chosen article should contain words that are new to the participants. These target words are either low-frequency words, or words whose meanings the participants may be familiar with but that pose predictable pronunciation challenges, such as unusual word stress and deceptive grapheme-phoneme correspondences.

Considering the aforementioned criteria, an introductory essay about climate change from a TED Study (https://www.ted.com/read/ted-studies/environmental-studies/introductory-essay) was chosen. Climate change is likely to be a topic of interest to the participants, especially given the more frequent and severe natural disasters occurring on every continent in recent years. As the majority of the participants are from arts and humanities or from social science backgrounds, they are unlikely to be familiar with some of the science-related terminology used in the article, where we can expect more pronunciation learning to happen.

To reduce the learning burden for the participants, some modifications were made to the article (See Appendix A for the modified version). First, the length of the article was reduced to <400 words. As part of this editing, a few low-frequency words that were located very close to each other were removed. However, one frequently mispronounced word (the verb *gauge*) was embedded to increase the number of target words.

The next consideration was the participants' text comprehension. One important predictor of text comprehensibility is the lexical profile of the text. I therefore ran the text through the VocabProfile tool in Lextutor (Cobb, n.d.). The results (see Table 1) indicate that the article consists of 362 tokens and that knowledge of the 4,000 (K4) most frequent word

families provides 96.1% lexical coverage. According to Laufer (2020), this should be sufficient for adequate text comprehension. Given the profile of our ESL learner participants, the majority of whom are graduate students and who had been immersed in an English-language environment for at least six months by the time of data collection, it is likely that they had receptive knowledge of most of the K1 to K4 word families of English. Glosses were provided in the margin of the text for the words that are beyond the K4 band. In addition to this written material, an audio-recording was created by an AI voice generator (https://voicemaker.in/), which is able to convert text into very human-like natural sounding voices. None of the participants reported they noticed that the audio was AI generated. To make it easier for our ESL participants to understand the audio in a single round of listening, I also reduced the playback speed by 15% with the help of the same website.

Table 1: *Lexical Profile of the Input*

Frequency level	Tokens	Cumulative tokens
1k level	71.8%	71.8%
2k level	11.7%	83.5%
3k level	11.0%	94.5%
4k level	1.6%	96.1%
5k level	1.6%	97.7%
6k level	0.6%	98.3%
7k level		
8k level	0.6%	98.9%
9k level	0.3%	99.2%
>10k level + off-list words	0.8%	100%

4.2.2 Target words

Sixteen target words were chosen from the article that were low-frequency words and/or words with predictable pronunciation issues (see Table 2). As noted, the selected pronunciation issues were vowels and/or word-stress placement. The decision to focus on these two elements as the target items was grounded in research highlighting their challenging nature and the crucial role they play for intelligibility.

In comparison to consonants, vowel sounds in English have more pronunciation variations, leading to a higher likelihood of incorrect guesses by L2 learners. Research by Munro and Derwing (1995) underscores that vowel errors can severely impact comprehensibility, more so than consonant errors. Therefore, targeting vowels can significantly help learners achieve more intelligible speech.

Word stress is also essential for effective communication. According to Field (2005), misplaced stress can make words unrecognizable to native listeners. This is supported by Gallego (1990), who found that communication breakdowns often occurred when nonnative English speaking teaching assistants made pronunciation errors, particularly with word stress. These findings underline the importance of word stress and suggest that mastering word stress can greatly improve L2 learners' intelligibility.

There are also studies that emphasize the importance of both vowels and word stress. (e.g., Derwing & Munro, 2005). Saito (2012) found that explicit instruction on these features not only improves pronunciation but also boosts learners' confidence and willingness to communicate. Additionally, Kang et al. (2010) and Saito et al. (2016) indicated that suprasegmental features like word stress and segmental features like vowels are closely linked to judgments of language proficiency and accentedness.

Most of the target words contained only one anticipated pronunciation challenge, whereas there were four words that required rating on two difficulties. These four words were distributed equally over the two halves of the article.

Table 2: *List of Target Words*

Word	IPA	Anticipated difficulties
drought	/draʊt/	Vowel [ou]
committee	/kəˈmɪt̬.i/	Word stress
obscure	/əbˈskjʊr/	Word stress & vowel [ure]
pivotal	/ˈpɪv.ə.t̬əl/	Word stress & vowels [i][o]
meteorological	/ˌmiː.t̬i.ɚ.əˈlɑː.dʒɪ.kəl/	Word stress
gauge	/geɪdʒ/	Vowel [au]
tenure	/ˈten.jɚ/	Vowel [ure]
equilibrium	/ˌiː.kwɪˈlɪb.ri.əm/	Word stress
contested	/kɑːnˈtestid/	Word stress
derisively	/dɪˈraɪ.sɪv.li/	Word stress & vowel [i] (the first one)
palpable	/ˈpæl.pə.bəl/	Vowel [a] (the first one)
politicized	/pəˈlɪt̬.ə.saɪzd/	Word stress & vowel [i] (the second one)
entrenched	/ɪnˈtrentʃt/	Word stress
unprecedented	/ʌnˈpres.ə.den.t̬ɪd/	Word stress
societal	/səˈsaɪ.ə.t̬əl/	Vowel/Diphthong [ie]
paralysis	/pəˈræl.ə.sɪs/	Word stress

4.3 Research Instruments – Interview and Questionnaire

To gain a more comprehensive understanding from participants' perspective, if individual factors, such as learning styles and language learning habits, affect the results of trial-and-error and retrieval procedures, a questionnaire (see Appendix B) that consists of thirty-six questions was included in the interview. The first questions aim at gaining a clearer picture of the participants' experience of the study; the following questions assess if participants habitually give precedence to aural input for language learning; the final part of the interview gauges the learning style of the participant, by using a questionnaire adapted from Cohen et al.'s (2009) *Learning Style Survey,* with a special interest in whether the participants are auditory learners or non-auditory learners.

4.4 Procedures

The study adopted a counterbalanced, within-participant design, where each participant experienced the two sequences of modality (reading + listening or listening + reading) but for different parts of the text. Participants were randomly assigned to one of two groups. Group A read the first half of the text silently and then listened to the audio-recording of the same part, and then listened to the second half of the text before reading that part silently. For Group B, the order was reversed. As shown in Table 3, all the participants completed the same tasks, but in different sequences according to this counterbalanced design.

The participants individually met the researcher in a quiet meeting room. Before the study, participants read the Letter of Information and then signed the Consent Form. The trial-and-error sequence first required students to read the text silently, then read it aloud (i.e. the 1st read-alouds), and then listen to the audio-recording. The aural-input-plus-retrieval sequence, by contrast, first had the students listen to the audio-recording of the text, after which they read the text silently and then aloud (i.e. the 1st read-alouds). The participants either listened to the passage or read it silently, but did not listen to the recording while reading the passage. To make sure that the participants processed the content of the text, they were told that content-related questions would follow the first time they read/ listened to it. There were six questions altogether, three for each of the

two parts of the text. After processing the whole text for the first time, the participants were asked to read aloud the entire text again (i.e. the 2nd read-alouds). In a final read-aloud activity (i.e. the 3rd read-alouds), the participants read the target words individually as the first step of the structured interview (see Appendix B). During all three read-aloud activities, I noted any hesitations in the participants' pronunciation of the target words, as these could indicate uncertainty or guessing, which might be useful for further data analysis.

Following the read-aloud, the structured interview continued with questions about how the participants had experienced the two learning procedures. This included whether they already knew any of the target words and their pronunciation prior to the experiment, whether they imagined the spoken forms while silently reading, whether they felt surprised by any of the spoken forms when they heard them in the audio-recording, whether they habitually give precedence to aural or written input for language learning, whether they sometimes listen to audio books in English, and so on. Asking the participants retrospectively if they already knew the target items and their pronunciation was preferred over administering a pre-test in this study, because administering a pre-test would itself constitute a trial-and-error event, and so it would compromise the comparison between the two treatment conditions (i.e., the independent variable of main interest). The interview also included questions to determine the learning style of the participant, adapted from Cohen et al.'s (2009) *Learning Style Survey*. The interview questions were initially designed as dichotomous, but participants were encouraged to elaborate on their responses and share additional details or reflections on the questions.

The length of the meeting ranged from forty minutes to one hour, depending on how long each participant took to complete the reading and comprehension questions, as well as how much they were willing to share during the interview segments.

Table 3 gives an overview of the steps or activities in the two learning conditions.

Table 3: *Trial-and-Error Procedures Versus Retrieval Procedures*

Step	**Group A**	**Group B**
Step 1	Silent reading of part 1 of the text + comprehension questions	Listening to part 1 of the text + comprehension questions
Step 2	Reading aloud part 1 of the text	Silent reading of part 1 of the text
Step 3	Listening to part 1 of the text	Reading aloud part 1 of the text
Short break		
Step 4	Listening to part 2 of the text + comprehension questions	Silent reading of part 2 of the text + comprehension questions
Step 5	Silent reading of part 2 of the text	Reading aloud part 2 of the text
Step 6	Reading aloud part 2 of the text	Listening to part 2 of the text
Short break		
Step 7	Reading aloud of the whole text	
Step 8	Structured Interview (Including reading aloud individual target words)	

Chapter 5 Data Coding and Analysis

This study collected data in the form of audio recordings, capturing participants' pronunciations of target words. These recordings were then converted into quantitative data for analysis. The way participants pronounced the target words was scored dichotomously: 1 point for a correct response and 0 point for an incorrect one, as determined by two human raters - the student researcher and the principal researcher.

The decision not to use a partially correct response option with 0.5 points was made to maintain clarity and consistency in the rating process. Introducing partial credit could complicate the scoring system, creating ambiguity about what qualifies as "partially correct," potentially leading to subjective judgments during assessment. What's more, the use of 0.5 points could limit the types of data analysis models that can be employed. For instance, while mixed-effects models can handle both continuous and categorical outcomes, introducing partial points could complicate interpretation, especially if model is designed for categorical data. Partial scores blur the distinction between categories, making it harder to differentiate the full impact of each variable. Partial credit might lead to unclear cutoffs or thresholds for meaningful learning improvements. This could make it harder to interpret learning outcomes and draw clear conclusions from the data.

Since the chosen target words posed specific pronunciation challenges, generating predictable errors, the raters assessed those issues exclusively. For example, challenges could be segmental, such as pronouncing a vowel (as /ei/ in *gauge*) or a diphthong (as /aɪ.ə/ in *societal*), or suprasegmental, notably stress placement (as in *equilibrium* and *paralysis*). The raters' task was specifically to assess the participant's performance regarding these anticipated issues. For twelve of the sixteen target words, only one pronunciation challenge required rating, while four words presented two difficulties (as shown in Table 2). When assessing these four words, participants received 1 point only if they correctly addressed both challenges; no half points were awarded for partially correct responses. If a participant substituted the target word with another (usually a more familiar word to them), it was coded as missing data.

As preparation, the raters listened to the sample audio-recording of the text at least five times, with the International Phonetic Alphabet (IPA) of each target word at hand, until they were familiar with the correct pronunciations. In addition, the raters discussed the anticipated difficulties of the target words to confirm each other's perception of the pronunciations. Both raters had experience in teaching students of L2 English, and so were familiar with accented speech, which can support comprehension (Gass & Varonis, 1984), particularly since one of the raters shared the L1 of the majority of the participants.

One rater initially scored all the responses, identifying forty-eight cases where they had lower confidence in their ratings. A second rater reviewed these cases and agreed with the majority of the ratings, with only two cases showing discrepancies. After further discussion, both raters reached a consensus on these two disagreements.

The interview data regarding participants' experience of the learning procedures, their learning habits and styles were coded. Responses of six yes-no questions were coded dichotomously, aimed to give a clearer picture of participants' prior knowledge of the target words, determine whether mental representations of the spoken form of the words were created during silent reading, and identify whether any spoken forms in the audio recordings captured their attention. These factors were crucial in assessing whether learning occurred and what processes drove this learning. Additionally, information about learning habits and styles was expected to help distinguish participants who stood out as auditory learners with a view to examining whether the treatment conditions have differential effects most noticeably for these participants.

A questionnaire (Appendix B), adapted from Cohen et al.'s (2009) *Learning Style Survey* were used to assess individual learning styles by evaluating responses to thirty items, with ten items each corresponding to auditory, visual, and tactile/kinesthetic preferences. To determine if a participant was an auditory learner, their total score for auditory-related items was compared to their scores for the other two learning styles. If the auditory score was the highest, the participant was categorized as an auditory learner.

The open-access platform *Jamovi* (Version 2.3.28, 2024) was used for the analyses. First, a generalized mixed effects model was conducted simply to examine if the differences in

error correction between the two learning conditions (trial-and-error versus retrieval practice) in the posttests were significant.

Next, to examine the role of other factors derived from the interviews, four stepwise logistic regression models were run, two for the trial-and-error condition responses (one for the 2nd read-aloud, and the other for the 3rd read-aloud) and two for the retrieval practice condition responses (one for each of the post-tests as well). The following fixed factors were initially included in the two models: a) participants' prior knowledge of the target words' meanings (yes or no), b) initial pronunciation accuracy, c) participants' recall of any surprise at the word sounds when listening to the audio-recording (yes or no), d) participants' learning styles (auditory versus non-auditory), e) participants' learning habits regarding listening to audiobooks (often or not), and f) type of pronunciation issue (vowel/diphthong versus stress placement). One additional initial factor was considered for the trial-and-error condition responses, notably participants' recall of imagining the sounds of the target words during silent reading. The random effects for both models were participants and items (target words). To obtain the best models, the non-significant factors were excluded in a stepwise fashion, starting with the highest p-values, until the models reached the smallest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

Chapter 6 Results

6.1 Scores

The descriptive statistics of the results (mean accuracy scores, standard deviations, minimum and maximum scores) in the three read-alouds are shown in Table 4.

Table 4: *Descriptive statistics for the three read-alouds*

	Read-Aloud #1		**Read-Aloud #2**		**Read-Aloud #3**	
Learning condition	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
Trial-and-error (*Listen-Second*)	0.555	0.498	0.610	0.488	0.610	0.488
Retrieval (*Listen-First*)	0.594	0.492	0.606	0.489	0.594	0.492

The summary of the error counts in the three read-aloud tasks is displayed in Table 5.

Table 5: *Summary of the error counts*

	Read-Aloud #1	**Read-Aloud #2**	**Read-Aloud #3**
Trial-and-error (*Listen-Second*)	173	151	156
Retrieval (*Listen-First*)	161	154	162

Of the total 800 responses (50 participants*16 target words), 334 (42.5%) were identified as pronunciation errors in the 1st read-aloud (initial errors). Based on the error counts, the Listen-First condition performed slightly better in the 1st read-aloud task, which indicates the usefulness of exposure to the spoken forms of words prior to reading, but the

difference compared to the Listen-Second condition was rather marginal. This suggests that listening to the audio-recording just once did not help a lot to prevent pronunciation errors.

In the 2nd read-aloud, it is the Listen-Second condition that showed more cases of improved pronunciation, with 22 errors corrected (173 – 151 = 22) compared to 7 cases of improvement in the Listen-First condition. This difference makes sense because learners in the Listen-Second condition listened to the audio-recording before their 2nd read-aloud and so were given an opportunity to learn the correct spoken word forms which they had mispronounced earlier, while there was no learning opportunity anymore after the 1st read-aloud for the Listen-First condition. It is, in fact, somewhat surprising that there were nonetheless 7 instances of improvement in the latter condition. In the 3rd (and final) read-aloud, the error rate increased slightly under both conditions, but with Listen-Second condition (i.e., the trial-and-error condition) producing slightly fewer errors than the Listen-First condition. In total, 17 (9.83%) of the 173 initial instances of incorrect pronunciation under the Listen-Second condition ended up correct in the final test, where the participants read each word in isolation.

To determine if any of the above differences were significant, mixed-effects modeling was used to analyze the data. The result (Table 3) shows that neither the learning condition nor the error correction made was a statistically significant factor in the 2nd and the 3rd read-aloud test performance.

Table 6: *Mixed-effects modeling for error correction in the 2nd and 3rd read-alouds (post-tests)*

				95% CI for *OR*			
Fixed Effects	Estimate	*SE*	*OR*	Lower	Upper	*z*	*p*
(Intercept)	0.566	0.354	1.76	0.879	3.52	1.597	0.110
Error Corrections in 2nd read-alouds (2 - 1)	0.202	0.124	1.22	0.961	1.56	1.635	0.102
Error Corrections in 3rd read-alouds (3 - 1)	0.184	0.123	1.20	0.945	1.53	1.499	0.134
Learning Condition	0.046	0.102	1.05	0.858	1.28	0.454	0.650
Error Corrections in 2nd read-alouds * Learning Condition	0.277	0.247	1.32	0.813	2.14	1.122	0.262
Error Corrections in 3rd read-alouds * Learning Condition	0.338	0.245	1.40	0.867	2.27	1.379	0.168

In addition, the participants' scores on the six comprehension questions ($M = 5.76$, $SD = 0.797$), evidenced that the participants listened/read for content and that the text was indeed comprehensible.

6.2 Incorporating the interview Data

To further understand the individual factors that influence participants' performance, responses of the interview questions were also analyzed, exclusively focusing on cases (*n*

= 288, 72.2% of the total cases) where participants reported not yet knowing the spoken form of the words before the experiment, as only these instances presented room for learning. If participants responded "no" to the interview question, "Did you already know their pronunciation prior to the experiment?" for certain target words, those cases were included in the 288 cases. For the trial-and-error condition, if participants reported knowing the spoken forms but pronounced them incorrectly during the first read-aloud, these instances were included in the 288 cases as well, since incorrect pronunciation before exposure suggests a lack of prior knowledge. This process helped minimize the potential inaccuracies associated with self-reports.

The analyses revealed that only one factor, initial pronunciation accuracy, was a statistically significant predictor of accurate pronunciation in the later read-loud tasks, and so only this factor was retained in the final logistic regression models. The other six initial fixed factors were excluded, as they were non-significant and did not improve the models: a) participants' prior knowledge of the target words' meanings (yes or no), b) participants' recall of any surprise at the word sounds when listening to the audio-recording (yes or no), c) participants' learning styles (auditory versus non-auditory), d) participants' learning habits regarding listening to audiobooks (often or not), e) type of pronunciation issue (vowel/diphthong versus stress placement), and f) participants' recall of imagining the sounds of the target words during silent reading.

Table 7 *Mixed-effects logistic regression for Listen-Second's pronunciation accuracy in the 2nd read-aloud*

				95% CI for *OR*			
Fixed Effects	Estimate	*SE*	*exp(B)*	Lower	Upper	*z*	*p*
(Intercept)	0.331	0.267	1.39	0.826	2.35	1.24	.214
1st read-aloud (1 - 0)	3.419	0.400	30.54	13.954	66.82	8.54	< .001

Random Effects	SD	Variance
Participants	0.000	0.000
Items	0.754	0.569

Notes. Number of Observations: 273.

R^2 marginal = 0.422; R^2 conditional = 0.508.

Table 8 *Mixed-effects logistic regression for Listen-Second's pronunciation accuracy in the 3rd read-aloud*

				95% CI for *OR*			
Fixed Effects	Estimate	*SE*	*exp(B)*	Lower	Upper	*z*	*p*
(Intercept)	0.240	0.220	1.27	0.826	1.96	1.09	.275
1st read-aloud (1 - 0)	2.614	0.338	13.65	7.034	26.49	7.73	< .001

Random Effects	SD	Variance
Participants	0.000	0.000
Items	0.600	0.360

Notes. Number of Observations: 280.

R^2 marginal = 0.309; R^2 conditional = 0.377.

Table 9 *Mixed-effects logistic regression for Listen-First's pronunciation accuracy in the 2nd read-aloud*

				95% CI for *OR*			
Fixed Effects	Estimate	*SE*	*exp(B)*	Lower	Upper	*z*	*p*
(Intercept)	-0.105	0.202	0.900	0.606	1.34	-0.521	.602
1st read-aloud (1 - 0)	3.651	0.362	38.502	18.928	78.32	10.076	< .001

Random Effects	SD	Variance
Participants	0.000	0.000
Items	0.382	0.146

Notes. Number of Observations: 276.

R^2 marginal = 0.491; R^2 conditional = 0.513.

Table 10 *Mixed-effects logistic regression for Listen-First's pronunciation accuracy in the 3rd read-aloud*

				95% CI for *OR*			
Fixed Effects	Estimate	*SE*	*exp(B)*	Lower	Upper	*z*	*p*
(Intercept)	0.240	0.220	1.27	0.826	1.96	1.09	.275
1st read-aloud (1 - 0)	2.614	0.338	13.65	7.034	26.49	7.73	< .001

Random Effects	SD	Variance
Participants	0.000	0.000
Items	0.600	0.360

Notes. Number of Observations: 280.

R^2 marginal = 0.309; R^2 conditional = 0.377.

For the Listen-Second condition, the results of both post-test models found a significant effect of initial pronunciation accuracy (i.e. the first read-aloud). Recall that the above analyses only considered the spoken forms of words that the participants reported not knowing beforehand. In the Listen-Second condition, participants had not yet experienced the learning procedure (i.e., listening to the audio-recording), and so any correct responses in the first read-aloud were accurate guesses or inferences, and all incorrect ones were inaccurate guesses. Results in Tables 7 and 8 show that accurate guesses were likely to predict better pronunciation performance in the post-tests, whereas failed ones were associated with less successful pronunciation production. Put differently, trial and error worked better if it involved no error from the start.

Results in Tables 9 and 10 indicate a similar trend existed in the Listen-First condition, though the interpretation of initial pronunciation accuracy was slightly more complex. Getting a score in the first read-aloud for the Listen-First condition could be due to either (1) successful retrieval of newly acquired pronunciation from the audio-recording, or (2) correct guessing. If the participants failed to obtain a point, it indicated that they neither retrieved the pronunciation successfully nor guessed correctly. While the information collected was insufficient to distinguish between retrieval and guessing in the Listen-First condition, correct pronunciation in the first read-aloud generally led to better pronunciation performance in the post-tests. The earlier descriptive statistics suggest that listening to the audio-recording first reduced the number of pronunciation errors in the first read-aloud task (relative to the trial-and-error condition), and so a certain amount of successful retrieval seems to have occurred, but it should not be overestimated since the difference in error rate between the two conditions was statistically non-significant. If, for

the majority of the target words, the Listen-First condition left no usable memories of the spoken forms, then the reading-aloud task that followed the listening activity may well have been experienced mostly as a trial-and-error activity as well.

Beyond the pre-designed interview questions, the following observations made during the read-aloud sessions may help to illustrate the complexity of the data set, and the limited number of statistically significant predictors. First, noticing seems to be an important factor. This finding was supported by the participants' explicit declarations as well as the researcher's observations of the participants' subtle behavior changes. The designed interview questions did not ask about participants' strategies or mental procedures for each error correction made; but three participants proactively reflected on and were excited to declare the connection between noticing and learning. Participant B3 made two error corrections: the one that was learned through noticing was stably produced correctly in both post-tests, whereas the other was only rated correct in the second read-aloud. Participant A15 recalled that she acquired the word stress of the target word from the audio, and consistently pronounced it with the correct word stress in both read-alouds.

Unlike Participant B3 and A15, whose word pronunciations were all learned under the trial-and-error condition, Participant B15 provided a comparison as she picked up new pronunciations from both the trial-and-error and the retrieval condition. For words that were learned from noticing under the trial-and-error condition, the success rate was higher than those learned in the retrieval condition. All trial-and-error words were produced successfully in the post-tests and even when they were mentioned later in the interviews, whereas the production of the retrieval words was more variable, with most of them only correctly produced one time in either post-test.

In addition to the direct evidence from participants' description, changes in their confidence levels between the first read-aloud and the two post-tests might also suggest that participants experienced noticing when learning. Five participants were observed to hesitate while pronouncing in total nine target words during in the first read-loud, which were later confirmed as correct. In the second and third read-aloud, they pronounced the same words confidently without hesitation. The initial hesitation might indicate that the participants doubted if their pronunciations were correct, or that they were trying to guess the correct pronunciation. These uncertainties likely prompted them to pay more attention to these words when they listened to the audio, and eventually their confidence in the post-tests suggest that they learned and confirmed the pronunciation of the target words.

The above examples involved participants who realized their lack of pronunciation knowledge for certain words and managed to adopt the correct pronunciation after noticing and listening just once. However, listening only once did not seem sufficient for forming solid memories for all participants. Some pronounced the words incorrectly during the two post-tests but recalled the correct pronunciation later during the interview, while others did the opposite, pronouncing the words correctly in the post-tests but not maintaining the correct pronunciation later in the interview. One participant's interview excerpt (translated from Mandarin) provided a possible reason for these varied learning effects: *"My attention went to those words that were somehow familiar, or where I had confidence in my guessing. It was similar to the concept of the Zone of Proximal Development. For those words that were completely unknown and perceived as difficult, I just let them slip."* In a nutshell, although participants had similar English proficiency

levels, their abilities to memorize word pronunciations from a single exposure to an audio-recording varied.

In a similar vein, participants appeared to generate varying levels of imagination regarding word pronunciation during silent reading. In the trial-and-error condition, whether participants imagined pronunciations in their mind during silent reading was the prerequisite to investigating if initial error-prone pronunciations could be easily remedied by implicit feedback (i.e. exposure to audio-recording). Out of the 50 participants, 24 (48%) reported that they did not imagine any sounds of the text. Even among those who did imagine sounds, the nature and extent of their imagination was not the same, as illustrated by the following interview excerpts (translated from Mandarin):

- *"I imagined when the word was familiar but I was not sure; I didn't imagine if I didn't know it at all."*
- *"During part 1 (TE) I didn't imagine, but in part 2 (R) I did."*
- *"I did imagine the sounds, but did not imagine every syllable."*
- *"I only imagined vowels and consonants, but not the word stress."*
- *"Yes, I did imagine but the sounds were not too clear, they were more like a background sound that I knew existed but were blurry."*

Regarding English language learning habits, participants seemed to place varying levels of importance on pronunciation training. Participants A4, B4, and B6 explicitly stated that they regarded speaking as the most important skill among the four language skills. Whenever they encountered new vocabulary during their daily English learning, they always looked up the words' pronunciations immediately. They believed that being certain about word pronunciations was the first step to eventually using them in oral

production. When I examined the scores of these three participants, they either corrected more errors than average (five versus 0.32 error corrections per participant in the 3rd read-aloud) or demonstrated better initial accuracy (thirteen/fourteen versus an average of 9.04 in the first read-aloud). Their proactive approach to pronunciation created more awareness of inconsistent spelling-sound correspondences, making them more cautious and attentive to the text to detect problematic words. Moreover, one participant provided another notable explanation (translated from Mandarin) – *"Phonetic intuition helps me guess pronunciation. Over the years, the more phonetic training I received and the more oral speeches I produced, the stronger my phonetic intuition developed. Compared to the earlier stages of my English learning journey, I now feel that I have a stronger phonetic intuition. I am able to find a comfortable way to pronounce new words, especially if I have chances to try them several times. Many of these guesses turned out correct."* Given the crucial role that initial accuracy played in predicting post-test performance, phonetic training might be a beneficial way to improve initial accuracy.

In fact, participants reported different levels of phonetics training experience. Instructions about the International Phonetic Alphabet (IPA) were mentioned by some, and a few of them mentioned phonics training too. The majority reported that they had never or seldom received any phonetics training, as it was not part of their formal school curriculum. Teachers might occasionally introduce vowels or consonants, but word stress was rarely emphasized. Without mastering systematic phonetic knowledge, such as the IPA, learners may struggle to perceive the nuances between the correct pronunciation they hear in samples of English speech and their own rendering of the same words, making it difficult to modify their pronunciation effectively. Only a few participants

declared that they had sufficient phonetic knowledge. Interestingly, these were the same ones who placed extra importance on pronunciation and said they always look up new words' pronunciations, when possible, mentioned in the earlier paragraph. Recall that these learners achieved better learning outcomes and/or initial accuracy; it could also be their phonetic knowledge/training that contributed to the better results.

In addition to the pre-selected target words, some participants picked up other words such as "fatalism," "dilemma," and the names "Thiel" and "Thatcher." If these words were proximal to the target words that were new to the participants, these words might have diverted participants' attention from the pre-selected target words. This shift in focus could have impacted their ability to correctly pronounce the target words, as their attention may have been divided.

Chapter 7 Discussion

In this chapter, the findings from the statistical analysis and the interview responses will be further discussed by linking them to the guiding theories.

7.1 RQ1: Can learners acquire word pronunciations from one incidental exposure to an audio-recording?

Incidental pronunciation acquisition from exposure to an audio-recording appears to be challenging even for upper-intermediate-level L2 learners. In this study, learners listened to an audio-recording with a focus on content (as they were informed to complete comprehension questions afterward) and were then asked to read the same text aloud. This "retrieval" condition was compared to a "trial-and-error" condition where the students first read the passage before listening to its audio-recording. In these first read-alouds, the retrieval condition showed only a slightly higher accuracy rate. Under the trial-and-error condition, students' pronunciation of some of the words improved after they had had a chance to listen to the recording, but most pronunciation errors persisted after listening. Therefore, we might conclude that listening to an audio-recording just once, and with a purpose other than learning pronunciation, is not very effective for acquiring word pronunciation.

However, the study did reveal that there is an association between correct initial pronunciation (of previously unfamiliar words) and accuracy in the post-test. This suggests that it is judicious to reduce the number of initial errors. Having students listen to the audio-recording first is one way to do so but a single exposure may not be sufficient. In the study, we slowed the audio to 85% of its original speed, and spaced out the target words, hoping to reduce difficulty and enhance pronunciation acquisition. However, even with these modifications, listening only once seems challenging for effective incidental learning. Repeated listening may be beneficial, as Webb and Nation (2017) suggest that repeated exposure frees up attentional resources, allowing learners to notice and focus on unknown words. In fact, it is common for many audio listeners to engage with the same material multiple times for different purposes (Wi, 2021).

Nevertheless, the spoken form of vocabulary is only one of the nine aspects of vocabulary knowledge (Nation, 2013), even if the repeated listening frees up learners' attentional resources from the content, it remains uncertain where and what learners direct their attention to. Therefore, paying attention to the spoken form of words and noticing knowledge gaps seem to be crucial for incidental pronunciation acquisition. The importance of noticing was evident from the participants' reflections in this study. Participants reported that the target words they successfully produced were those for which they identified pronunciation knowledge gaps and to which they paid extra attention during listening activities.

To increase the likelihood of incidental pronunciation learning, it may be important for learners to be more alert towards unexpected pronunciations. Incorporating more opportunities for repetition and focused listening may enhance acquisition. For example, repeated exposure to the same audio material for different purposes can aid in this process. Focused listening could be done as prior practice that includes structured activities that prompt learners to listen for specific pronunciation features, improving their ability to perceive and encode pronunciation features. This heightened sensitivity will help them notice unexpected pronunciations more effectively. Given the nature of incidental learning, expectations regarding the number of word pronunciations acquired from a single material should be managed. The goal should be gradual improvement and increased awareness, rather than immediate mastery of numerous new pronunciations.

In summary, while one incidental exposure to an audio recording can provide some benefits in word pronunciation acquisition, it seems insufficient for substantial learning. To improve the effectiveness, repeated listening and prior focused listening training, along with strategies to reduce initial errors and to encourage active noticing, might be helpful.

7.2 RQ2: Is the trial-and-error or retrieval procedure the more effective approach for learning the pronunciation of new English words?

The error count summary illustrates that the retrieval condition produced fewer initial errors than the trial-and-error condition. Further improvement was unlikely, because this condition did not include additional use of audio recording. As expected, the trial-and-error condition prompted more initial errors, but a few of these were rectified after exposure to the audio recording. None of the differences in error counts in 2nd and 3rd read-aloud tasks were statistically significant, indicating the final learning outcomes were similar under both conditions. This finding suggests that the two approaches eventually converged to produce comparable outcomes.

The fewer initial errors in the retrieval condition may indicate the benefit of providing error-free input before reading the text aloud, which is supported by research (e.g., Wang et al., 2022; Strong & Boers, 2019) that suggests error-free learning processes can avoid interference from wrong guesses that are likely to linger in memory. On the other hand, the reduction of errors in the trial-and-error condition may be linked to the Output Hypothesis. According to this hypothesis, producing language (output) helps learners to recognize gaps in their knowledge, and to fill these gaps thanks to feedback or when they encounter the required language elements or features later. In the Listen-Second condition, participants first attempted to pronounce the words, which may have heightened their awareness that they were not sure of these spoken forms. When they then listened to the audio recording, they could compare their output with the correct pronunciation, facilitating a learning process. This learning effect appeared confined to only a small number of words, however.

7.3 RQ3: What are the individual factors that might affect the learning outcomes, according to the participants' retrospective verbal reports?

Early in the design stage of this study, we considered the significant role that individual factors could have on the learning outcomes. Therefore, during the interviews, questions were asked about a) participants' prior knowledge of the target words' meanings and pronunciations, b) participants' recall of any surprise at the word sounds when listening to the audio-recording, c) participants' learning styles (auditory versus non-auditory), d) participants' learning habits regarding listening to audiobooks (often or not), and e) participants' recall of imagining the sounds of the target words during silent reading. However, none of these factors were significant predictors of learning outcomes. A more qualitative analysis of the interview excerpts nonetheless suggested that the following characteristics might have had some impact: 1) levels of imagination regarding word pronunciation during silent reading, 2) thoughts on the importance of pronunciation, and 3) phonetic knowledge and training experience.

The level of imagination is important because incorrect imagination during the trial-and-error condition could potentially disadvantage learning outcomes, a hypothesis we aimed to investigate. Almost half of the participants reported not imagining any sounds, and for those who did imagine spoken word forms, the extent and number of words they imagined varied. Since we did not delve deeply into this question during the interviews, accurate measurements of the level of imagination were missing. Including this in future studies might prove useful.

Despite most participants being from similar programs, they had varying goals of English language learning. Learners who placed more emphasis on pronunciation and actively sought correct pronunciations, rather than other aspects of vocabulary knowledge, were more likely to acquire pronunciation from listening.

Learners' prior phonetic knowledge and training experiences also seemed to play a role. Participants with more extensive phonetic training reported better initial accuracy and

greater improvement in pronunciation. This suggests that a solid foundation in phonetics might enhance the ability to perceive and produce correct pronunciations.

Although data analyses indicated that learning styles did not significantly predict the learning outcomes in this study, improving the measurement method might yield different results. The instrument we used, adapted items from Cohen et al.'s (2009) *Learning Style Survey*, might not have been optimal for detecting the role of learning style in this study. None of the survey items were closely related to L2 processing, let alone L2 processing during reading. This may explain why some participants expressed uncertainty about the survey. They were unsure whether to answer based on their first language or second language (English) learning habits, as they noticed differences between them. What's more, many of the survey items focused on visual learning styles that prioritized the use of visual cues, very few distinguishing "print-oriented" learners who favored text. Additionally, some participants had very close scores in auditory and other learning styles. For instance, Participant B21 scored 22 in visual learning and 23 in auditory learning. The marginal difference between these scores complicates categorizing the participant strictly as a visual or auditory learner. However, following the survey instructions, this participant was categorized as an auditory learner.

Chapter 8 Implications

The findings of this study have implications for English word pronunciation acquisition for adult second language learners. As far as we know, these are the first experimental data to provide insights into the effectiveness of specific learning sequences – should learners listen first or read first? The following implications are drawn from the study:

1. **Input-Output-Input sequence (listen-read-listen again) may reap the benefits of the two approaches and lead to better learning outcomes.** Listening first can be beneficial to reduce initial errors. This initial exposure helps learners form a correct mental representation of the word's pronunciation, which can prevent the establishment of incorrect phonological representations. Next, complementing listening with opportunities for active production is recommended. Similar suggestions were made by Nguyen and Boers (2019). During the output task, learners can reflect on their original knowledge, realize the differences from the correct pronunciation, and prepare to adjust, when the initial attempt is not correct. For those correct attempts, learners' experimenting with the newly acquired pronunciation during the output task aids in entrenching these in memory. Lastly, another exposure to the same input helps solidify the new or improved pronunciation.

2. **During silent reading, L2 learners may be encouraged to look up and listen to the correct pronunciation of new words, to avoid creating inaccurate phonological representations.**
   This proposal to present learners with both the written and spoken forms of new L2 words simultaneously is advocated by Uchihara et al. (2022), with a view to facilitating the development of form-meaning connections and accurate pronunciation. Although written resources remain the most widely available and easily accessible in many educational settings, it is worth adopting reading materials with audio support; to use language learning applications that co-present both written and spoken forms of words. For example, instead of using a silent flashcard, use one that integrates pronunciation of words; turn on the English

captions when watching TV shows, movies, and video clips; use digital dictionaries that can play the word pronunciations, especially for learners who are unfamiliar with a phonetic script such as IPA.

3. **Reading aloud activities**, as in reading a story to someone, or reading out a speech script, may be useful to make readers aware of the gaps in their knowledge of the spoken form of words. Recognizing learners' pronunciation problems during such activities provides an opportunity for instructors to provide corrective feedback.

4. **Reading while listening** to an audio-recording of the text could help learners notice unexpected correspondences between the written and spoken forms of words more directly. Processing the written and spoken forms simultaneously instead of consecutively may make it easier to compare the spoken and written forms. Watching audio-visual materials (video) with closed captions is a straightforward way of applying this simultaneous processing.

5. **Integrating phonetic training into the curriculum could enhance learners' ability to perceive and produce accurate pronunciation.** As noted, participants with more phonetic training showed better initial accuracy and learning outcomes. Incorporating systematic phonetic instruction can provide learners with the tools to decode and reproduce correct pronunciation more effectively. This training may help learners develop a stronger phonetic intuition, enabling them to make more accurate guesses about the pronunciation of new words and reinforcing their ability to correct errors.

Chapter 9 Limitations and Future Research

This study has several limitations that need to be acknowledged. First, the relatively small sample size of 50 participants may impact the likelihood of finding statistically significant outcomes. Although a within-participant design does not require as large a sample as a between-participant design, the total number remains modest. What's more, this study only included upper-intermediate English learners, which may limit the generalizability of the findings. Results might differ with lower-level or higher-level learners. For instance, lower-level learners may struggle more with pronunciation, making it harder for them to correct mistakes, whereas higher-level learners may already know most of the target words, resulting in fewer mistakes to correct. Additionally, it would be interesting to investigate whether learners with varying levels of proficiency within the upper-intermediate group showed different patterns of learning gains. It is possible that learners who made fewer initial mistakes, due to their stronger prior knowledge, might have benefited more from the listening phase, aligning with the Hawthorne effect (Sedgwick & Greenwood, 2015), where those with greater initial knowledge ("the rich") improve faster than those with less knowledge ("the poor").

The study utilized only 16 target words, which is a limited number. Embedding more unfamiliar words in the text would likely compromise the ecological validity of the materials as it would require excessive use of glossing to maintain the comprehensibility of the text for participants.

Another limitation concerns the measurement of prior knowledge of the pronunciation of the target words in the listening-first condition. It cannot be ruled out that participants who pronounce target words correctly during their read-aloud task already knew their pronunciation even before listening to the audio-recording of the text, and so the correct pronunciation need not be attributed to the immediately preceding aural input. Asking these participants retrospectively about this prior knowledge is arguably not ideal because their responses may not be entirely reliable. As explained previously, however, this was nonetheless preferred over administering a pretest that asks the participants in this treatment condition to pronounce the target words because this would itself constitute a

trial-and-error activity and would thus compromise the intended contrast between the two conditions. An alternative could be to use pseudowords as targets for learning, and determine their "correct" pronunciation, but this was considered difficult to justify to students who expect that participating in a study about L2 learning will benefit their own acquisition of English.

Another limitation is that we do not know to what extent some of the correct pronunciations in the post-tests were lucky guesses. It could have been useful in the interview to ask the participants for confidence ratings when they re-read the target words (how confident are you about your pronunciation of this word on a scale from 0 to 5). Conceivably, a difference between the two conditions may not emerge in actual error rates, but nonetheless a difference may be noticeable at the level of confidence.

Additionally, the study did not control for or measure the time each learner spent on each target word during silent reading. The longer the learner focuses on one word, the more processing effort they might invest in the word. Unlike listening to an audio recording, the pace of which is controlled, and where every learner listens to each target word once, learners can voluntarily allocate time differently to words during silent reading. This could be a confounding factor, but it may also be argued that it is inherent to the distinction between silent reading and real-time listening.

Future studies might benefit from employing eye-tracking technology to measure how much attention a learner allocates to each target word (Pellicer-Sánchez & Perez, 2024), and see if this is associated with the learner's creating a mental representation also of the spoken form of the words. Follow this as the next step, it would be also interesting to look at if the words' lexical properties, such as word length, word frequency, familiarity, and predictability, influence attention and then formation of mental representation. In this study, we attempted to control these factors by using a counterbalanced design.

References

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the real world: Essays illustrating fundamental contributions to society*, *2*(59-68).

Boers, F. (2021). *Evaluating second language vocabulary and grammar instruction: A synthesis of the research on teaching words, phrases, and patterns*. Routledge.

Caravolas, M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective. *European Psychologist*, *9*(1), 3-14.

Cerezo, L., Caras, A., & Leow, R. P. (2016). The effectiveness of guided induction versus deductive instruction on the development of complex Spanish gustar structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition*, *38*(2), 265-291.

Cobb, T. (n.d.). Vocabprofile [Computer program]. Retrieved October 4, 2023, from https://www.lextutor.ca/vp/comp/

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL quarterly*, *39*(3), 379-397.

Derwing, T. M., Munro, M. J., & Thomson, R. I. (Eds.). (2022). *The Routledge Handbook of Second Language Acquisition and Speaking.* London: Routledge.

Derwing, T., Munro, M., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning, 48*, 393-410.

Dörnyei, Z. (2005). *The Psychology of the Language Learner: Individual Differences in Second Language Acquisition*. Lawrence Erlbaum Associates.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner: Revisited*.

Routledge.

Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL quarterly*, *39*(3), 399-423.

Gallego, J. C. (1990). The Intelligibility of Three Nonnative English-Speaking Teaching Assistants: An Analysis of. *Issues in Applied Linguistics*, *1*, 2.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding*? Memory & Cognition, 40*, 505–513.

Griffiths, C. (2012). Learning styles: Traversing the quagmire. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for Language Learning: Insights from Research, Theory and Practice* (pp. 151–168). Palgrave Macmillan.

Haight, C. E., Herron, C., & Cole, S. P. (2007). The effects of deductive and guided inductive instructional approaches on the learning of grammar in the elementary foreign language college classroom. *Foreign language annals*, *40*(2), 288-310.

Herron, C., & Tomasello, M. (1992). Acquiring grammatical structures by guided induction. *The French Review*, *65*(5), 708-718.

Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition, 40*, 514–527.

Hulstijn, J. H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*, 539–558.

Kang, O., Rubin, D. O. N., & Pickering, L. (2010). Suprasegmental measures of accentedness and judgments of language learner proficiency in oral English. *The Modern Language Journal*, *94*(4), 554-566.

Karpicke, J. D., Lehman, M., & Aue, W. R. (2014). Retrieval-based learning: An episodic context account. *Psychology of Learning and Motivation, 61*, 237–284.

Kessler, B. (2009). Statistical learning of conditional orthographic correspondences. *Writing Systems Research*, *1*(1), 19-34. DOI: 10.1093/wsr/wsp004

Kornell, N., Hays, M. J. & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 989–998.

Laufer, B. (2020). Lexical coverages, inferencing unknown words and reading comprehension: How are they related? *TESOL Quarterly, 54*(4), 1076–1085. https://doi.org/ 10.1002/tesq.3004.

Levis, J. M. & Sonsaat, S. (2020). Publication venues for L2 pronunciation research. *Journal of Second Language Pronunciation, 6*, 1–11.

Li, M. (2023). *Learning verb-noun collocations through multiple-choice exercises: Do distractors benefit or hinder later recall?* (MA Dissertation, University of Western Ontario).

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*(1), 465–489. https://doi.org/10.1146/annurev-psych-010416-044022

Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners? *The Canadian Modern Language Review, 63*(1), 127–147. DOI:10.3138/cmlr.63.1.127.

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language learning*, *45*(1), 73-97.

Nation, I. S. P. (2013). *Learning vocabulary in another language (2nd ed.).* Cambridge University Press.

Pellicer-Sánchez, A., & Perez, M. M. (2024). Eye-tracking in vocabulary research: Introduction to the special issue. *Research Methods in Applied Linguistics*, *3*(1), 100095.

Reid, J. M. (1995). *Learning styles in the ESL/EFL classroom*. Heinle & Heinle Publishers.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in second language acquisition*, *18*(1), 27-67.

Roediger III, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*, 20–27.

Rosa, R., & O'Neill, M. D. (1999). Explicitness, intake, and the issue of awareness. *Studies in Second Language Acquisition, 21*, 511-556.

Saalfeld, A. (2011). Acquisition of L2 phonology in advanced learners: Does instruction make a difference? In. J. Levis & K. LeVelle (Eds.). *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference,* Sept. 2010. (pp. 144-152), Ames, IA: Iowa State University.

Saito, K. (2012). Effects of Instruction on L2 Pronunciation Development: A Synthesis of 15 Quasi-Experimental Intervention Studies. *TESOL Quarterly*, *46*(4), 842–854. http://www.jstor.org/stable/43267893

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.

Saiz Aja, A. M. (2022). *Relationship among perceptual learning styles, productive vocabulary, and textbooks in English as a foreign language learning* (Doctoral dissertation, Universidad de La Rioja).

Schmidt, R. W. (1990). The role of consciousness in second language learning1. *Applied linguistics*, *11*(2), 129-158.

Sedgwick, P., & Greenwood, N. (2015). Understanding the Hawthorne effect. *Bmj*, *351*.

Shaffer, C. (1989). A comparison of inductive and deductive approaches to teaching foreign languages. *Modern Language Journal, 73*, 395-403.

Smith, M. A., Roediger, H. L., & Kapicke, J. D. (2013). Covert retrieval practice benefits retention as much as overt retrieval practice. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1712–1725.

Strong, B., & Boers, F. (2019). The error in trial and error: Exercises on phrasal verbs. *TESOL Quarterly, 53,* 289–319.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, *36*(3), 326-344.

Uchihara, T., & Harada, T. (2018). Roles of vocabulary knowledge for success in English-medium instruction: Self-perceptions and academic outcomes of Japanese undergraduates. *TESOL Quarterly, 52*(3), 564–587. 10.1002/tesq.453.

Uchihara, T. (2022). Is it possible to measure word-level comprehensibility and accentedness as independent constructs of pronunciation knowledge? *Research Methods in Applied Linguistics*, *1*(2), 100011.

Vogel, S., Herron, C., Cole, S. P., & York, H. (2011). Effectiveness of a guided inductive versus a deductive approach on the learning of grammar in the intermediate-level college French classroom. *Foreign language annals*, *44*(2), 353-380.

Wang, X., Boers, F., & Warren, P. (2022). Prompting language learners to guess the meaning of idioms: do wrong guesses linger? *Language Awareness*, 1-16.

Yu, X., & Boers, F. (2023). Inferring the meaning of idioms: Does accuracy matter for retention in memory? *RELC Journal*.

Zwier, L. J., & Boers, F. (2023). *English L2 Vocabulary Learning and Teaching: Concepts, Principles, and Pedagogy*. Taylor & Francis.

Uchihara, T., Webb, S., Saito, K., & Trofimovich, P. (2022). Does mode of input affect how second language learners create form-meaning connections and pronounce second language words? *The Modern Language Journal*, 106(2), 351–370. https://onlinelibrary.wiley.com/doi/full/10.1111/modl.12775

Wi, I. (2021). *Learning New Words from TED Talks: Strategic use of L1 subtitles and L2 captions*. [Master's thesis, Western University]. Electronic Thesis and Dissertation Repository. 8283. https://ir.lib.uwo.ca/etd/8283

Appendices

Appendix A

Text

Part 1:

One of the worst natural disasters in the U.S. was the drought over the summer of 1988. Thousands of people died during the intense heat waves.

drought: a long period of dry weather.

It was against this backdrop that scientist James Thiel delivered his predictions about global warming to the Energy and Natural Resources Committee. The next day, The New York Times ran a headline that read "Global Warming Has Begun." This transformed climate change from an obscure topic to one that people began to discuss over dinner.

That same pivotal year, the World Meteorological Association hosted the first conference to gauge the risks of climate change. And it was when British Prime Minister Margaret Thatcher, in her final tenure, recognized the problem: "For generations, we have assumed that mankind would have no impact on the equilibrium of the world's atmosphere," she said. "But it now appears that we have begun a massive experiment that puts our planet in danger."

Meteorological: relating to weather forecasting atmospheric sciences

tenure: The holding of an office

equilibrium: a state of physical balance

Part 2:

Climate change is still hotly contested today, several decades later, with climate deniers derisively rejecting the evidence regardless how compelling it is. Tensions remain palpable, as when nearly 800 NGO representatives walked out of the 2013 international climate negotiations in Poland.

derisively: in a manner expressing mocking

palpable: so intense as if it's touchable

How has climate change become so politicized? It requires us to tackle difficult dilemmas, like how the least developed nations will cope with its effects and who should help them financially. It highlights serious economic challenges like how to reckon with entrenched carbon-based industries, such as the big petrol companies whose shareholders make huge profits. Changing from carbon-based energy to green energy calls for global collaboration on an unprecedented scale.

entrenched: difficult or unlikely to change

Also, the fast pace of climate change can prompt societal fatalism, so that large groups of people no longer believe that any steps can be taken to stop it. Because of this paralysis, even just slowing down global warming may prove impossible. The topic remained newsworthy over the rest of that pivotal year. The World Meteorological Association hosted a conference called "Our Changing Atmosphere." This was one of the earliest international conferences to gauge the risks of climate change.

fatalism: the belief that all events are predetermined and therefore inevitable.

paralysis: inability to act

Comprehension Questions

Part 1:

1. Why did the summer of 1988 become a crucial moment for climate change awareness?

a. Due to a series of damaging hurricanes
b. Because of the prolonged dry weather and intense heat waves
c. Owing to intense debates by scientists and politicians

2. What was the reaction of the public to the headline "Global Warming Has Begun" in *The New York Times*?

a. Increased discussion
b. Lack of interest
c. Denial of climate change

3. How did the recognition of climate change by British Prime Minister Margaret Thatcher differ from the previous assumption about mankind's impact on the atmosphere?

a. She devalued the significance of climate change
b. She blamed natural processes for climate change
c. She acknowledged a potential danger from human activities

Part 2:

1. **Why did nearly 800 NGO representatives walk out of the 2013 international climate negotiations in Poland?**

a. Due to an unexpected breakthrough in climate change solutions
b. Because of tensions and disagreements surrounding climate change issues
c. To express support for the carbon-based energy industry

2. **What is emphasized as a necessity for addressing the economic challenges posed by climate change?**

a. Increased profitability of carbon-based industries
b. Collaboration on an unprecedented scale for green energy initiatives
c. Reduction of financial assistance to least developed nations

3. **What challenge arises from the belief that large groups of people no longer think any steps can be taken to address the fast pace of climate change?**

a. Accelerated transition to green energy
b. Increased financial support for least developed nations
c. Inability in taking steps to slow down global warming

Appendix B

Structured Interview Questions

1. Did you already know the meaning of any of the target words prior to the experiment?
2. Did you already know their pronunciation prior to the experiment?
3. Did you imagine the spoken forms of the target words while reading silently?
4. Did you feel surprised by any of the spoken forms when you heard them in the audio-recording?
5. Do you usually tend to learn from aural first or written input first for language learning?
6. Do you sometimes listen to audio books in English?

learning style questions (adapted from:
Cohen, A. D., Oxford, R. L., & Chi, J. C. (2009). *Learning Style Survey: Assessing your own learning styles*. Center for Advanced Research on Language Acquisition, University of Minnesota.)

Learning Style Survey: Assessing your own learning styles
For each item, circle your response:
0 = Never
1 = Rarely
2 = Sometimes
3 = Often
4 = Always

1. I remember something better if I write it down. 0 1 2 3 4
2. I take detailed notes during lectures. 0 1 2 3 4
3. When I listen, I visualize pictures, numbers, or words in my head. 0 1 2 3 4
4. I prefer to learn with TV or video rather than other media. 0 1 2 3 4
5. I use color-coding to help me as I learn or work. 0 1 2 3 4
6. I need written directions for tasks. 0 1 2 3 4
7. I have to look at people to understand what they say. 0 1 2 3 4
8. I understand lectures better when professors write on the board. 0 1 2 3 4
9. Charts, diagrams, and maps help me understand what someone says. 0 1 2 3 4
10. I remember peoples' faces but not their names. 0 1 2 3 4
A – Total:

11. I remember things better if I discuss them with someone. 0 1 2 3 4
12. I prefer to learn by listening to a lecture rather than reading. 0 1 2 3 4
13. I need oral directions for a task. 0 1 2 3 4
14. Background sound helps me think. 0 1 2 3 4
15. I like to listen to music when I study or work. 0 1 2 3 4
16. I can understand what people say even when I cannot see them. 0 1 2 3 4

17. I remember peoples' names but not their faces. 0 1 2 3 4
18. I easily remember jokes that I hear. 0 1 2 3 4
19. I can identify people by their voices (e.g., on the phone). 0 1 2 3 4
20. When I turn on the TV, I listen to the sound more than I watch the screen. 0 1 2 3 4

B – Total:

21. I'd rather start to do things, rather than pay attention to directions. 0 1 2 3 4
22. I need frequent breaks when I work or study. 0 1 2 3 4
23. I need to eat something when I read or study. 0 1 2 3 4
24. If I have a choice between sitting and standing, I'd rather stand. 0 1 2 3 4
25. I get nervous when I sit still too long. 0 1 2 3 4
26. I think better when I move around (e.g., pacing or tapping my feet). 0 1 2 3 4
27. I play with or bite on my pens during lectures. 0 1 2 3 4
28. Manipulating objects helps me to remember what someone says. 0 1 2 3 4
29. I move my hands when I speak. 0 1 2 3 4
30. I draw lots of pictures (doodles) in my notebook during lectures. 0 1 2 3 4

C – Total:

Curriculum Vitae

Name:	Liwen Situ
Post-secondary Education and Degrees:	Western University London, Ontario, Canada 2013-2016 B.A. Western University London, Ontario, Canada 2021-2022 M.Ed. Western University London, Ontario, Canada 2022-2024 M.A.
Honours and Awards:	Western Graduate Research Scholarship 2022-2024
Related Work Experience	Research Assistant The University of Western Ontario 2022

Publications:

Situ, L. (2024, July). *Listening First? Reconsidering the Approach to English Word Pronunciation Learning from Textual Input.* Presented (Poster) at EuroSLA 33, Montpellier, France.

Situ, L. (2024, April). *Listening First? Reconsidering the Approach to English Word Pronunciation Learning from Textual Input.* Presented (Poster) at 2024 Faculty Research Partner Day, Western University, Ontario, Canada.

Han, M., Heidarian, Z., Liu, C., Liu, H., Marion, J., Seabrook, F., **Situ, L.,** Wu, L., Taylor, S. (2024, April). *Reflections on the role of identity texts in English Language Teaching (ELT).* Co-presented online at TESOL 2024 International Convention & Expo, Tampa, Florida, USA.

Situ, L. (2024, March). *Listening First? Reconsidering the Approach to English Word Pronunciation Learning from Textual Input.* Presented (Poster) at Robert Macmillan Symposium in Education, Western University, Ontario, Canada.

Situ, L. (2024, March). *Learning English Word Pronunciation through Text: Should We Listen First?*
A Comparison of Trial-and-error and Retrieval Practice Sequences. Presented at TESOL 2024 International Convention & Expo - Graduate Student Forum, Tampa, Florida, USA.

Situ, L. (2024, March). *Exploring the Optimal Sequence for English Word Pronunciation Acquisition through Text: Is Prioritizing Listening the Key?* Presented online at Rosa Bruno Jofre Symposium in Education, Queen's University, Kingston, Ontario, Canada.

Han, M., Heidarian, Z., Liu, C., Liu, H., Cerqueira, F. M., Marion, J., Seabrook, A., **Situ, L.,** Wu, L., Taylor, S. (2023, May). *Critically Reflecting on Identity Texts: Equity-Based Praxes and Teacher Education.* Co-presented at the 4th Symposium of Southern Ontario Universities, Mississauga, Ontario, Canada.