EXPLICIT RULES OR IMPLICIT IMITATION: A COMPARATIVE STUDY OF TWO APPROACHES TO TEACHING ENGLISH PROSODY TO CHINESE EFL ADULT LEARNERS

Xiaodan Zhang

Explicit Rules or Implicit Imitation: A Comparative Study of Two Approaches to Teaching English Prosody to Chinese Adult EFL Learners

Xiaodan Zhang

TESI DOCTORAL – TESIS DOCTORAL – DOCTORAL THESIS

2024

Explicit Rules or Implicit Imitation: A Comparative Study of Two Approaches to Teaching English Prosody to Chinese Adult EFL Learners

Xiaodan Zhang

Ph.D. Thesis

Supervised by Dr. Joaquín Romero Gallego

Submitted to the

Department of English and German Studies

UNIVERSITAT ROVIRA I VIRGILI

2024

To my dearest mum and dad

Every triumph, every success, and every joy in my life is a reflection of the love, and

sacrifices you've showered upon me. In gratitude and with heartfelt appreciation, I

dedicate all my achievements to you.

I shall carry your love with me in every future step.

Thank you for being my inspiration and my greatest source of strength.

UNIVERSITAT ROVIRA i VIRGILI

FAIG CONSTAR que aquest treball, titulat "Explicit Rules or Implicit Imitation: A Comparative Study of Two Approaches to Teaching English Prosody to Chinese EFL Adult Learners", que presenta Xiaodan Zhang per a l'obtenció del títol de Doctor, ha estat realitzat sota la meva direcció al Departament d'Estudis Anglesos i Alemanys d'aquesta universitat.

HAGO CONSTAR que el presente trabajo, titulado "Explicit Rules or Implicit Imitation: A Comparative Study of Two Approaches to Teaching English Prosody to Chinese EFL Adult Learners", que presenta Xiaodan Zhang para la obtención del título de Doctor, ha sido realizado bajo mi dirección en el Departamento d'Estudis Anglesos i Alemanys de esta universidad.

I STATE that the present study, entitled "Explicit Rules or Implicit Imitation: A Comparative Study of Two Approaches to Teaching English Prosody to Chinese EFL Adult Learners", presented by Xiaodan Zhang for the award of the degree of Doctor, has been carried out under my supervision at the Department d'Estudis Anglesos i Alemanys of this university.

Joaquín Romero

Digitally signed by Joaquín Romero
Date: 2024.01.23 13:19:48 +01'00'

Dr. Joaquín Romero Gallego

Doctoral Dissertation Supervisor

Tarragona, January 23, 2024

v

Declaration

I, Xiaodan Zhang, hereby declare that this dissertation is entirely my own work, carried out at Rovira i Virgili University for the degree of Doctor of Philosophy, and the contents of this dissertation are original and have not been submitted in whole or in part for a degree or qualification at any other university. Where other sources of information have been used, they have been acknowledged.

Some parts of this dissertation have been accepted for publication in Zhang, X., & Romero, J. (2023). Explicit rules or implicit imitation: A comparative study of two approaches to teaching English prosody. In R. Skarnitzl, & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 1613–1617). Guarant International. Another is accepted to publish on The Journal of Monolingual and Bilingual Speech (JMBS).

Correlated studies can be found at: i) Zhang, X., & Romero, J. (2021). Investigation into the correlation between perception and production of English sentence stress by Chinese EFL learners. In Ruiz-Cecilia, R. & Guijarro-Ojeda, J. R. (Eds.). *Investigación e innovación en lengua extranjera: Una perspectiva global (pp. 879-905)*. Tirant lo Blanch. ii) Zhang, X., & Romero, J. (2021). Investigating the influence of sentence stress perception on foreign accent in utterances. In *Proc. 3rd International Symposium on Applied Phonetics* (*ISAPh 2021*) (pp. 76-79).

Xiaodan Zhang

Tarragona, January 23, 2024

Acknowledgements

First and foremost, my deepest gratitude goes to my supervisor, Dr. Joaquín Romero Gallego, whose guidance has been the cornerstone of my doctoral studies. Undertaking doctoral research is arduous, but Joaquín's unwavering encouragement, vast knowledge, boundless patience, and motivational spirit have not only made my doctoral journey fruitful but also immensely enjoyable. In him, I've found not just a supervisor but a life mentor, and for that, I am sincerely thankful from the bottom of my heart.

I would also like to express my sincere gratitude to the Department of English and German Studies at Rovira i Virgili University, a place where I have experienced substantial academic, professional, and personal growth. This division not only witnessed the completion of my master's and PhD studies but also provided me with the invaluable opportunity to embark on my teaching career as part of its team.

My special appreciation goes to Leticia Quesada, whose generosity in sharing resources and cheerful presence in the lab made my experience even more delightful. I feel fortunate to have such a wonderful colleague, and the memorable moments we shared at the ICPhS conference in Prague are deeply treasured.

To all my cherished fellows in our phonetics lab, you have added great joy to every minute I spend there. I am fortunate to have commenced this long journey with such a congenial group. Particularly, a heartfelt acknowledgement is extended to Leonardo Oliveira, who has always been so generous and cheerful in sharing his knowledge and resolving my doubts.

Lastly, to my soulmate—my husband Yuan—and our adorable boy Lucas, you are the sunshine that makes each of my days beautiful. I cannot express how much I love you both; thank you for being my rock and source of joy!

Abstract

This study aims to explore the effect of explicit and implicit instructions on improving English pronunciation as far as stress production is concerned. 28 tertiary-level Chinese EFL learners were divided equally into two groups and attended a pronunciation training course of eight 30-minute sessions, instructed using the two approaches in question respectively. The pretest and the posttest, involving a controlled reading task and a spontaneous speaking task, were conducted before and after the experiment.

Participants' recordings were analyzed acoustically on PRAAT. Readings of intensity peaks, F0 averages and duration for each syllable at word-level and for each word at sentence-level were obtained. These values were then processed in order to obtain the difference in amplitude and F0, together with the duration ratios, between stressed and unstressed syllables/words. The statistical results generated from a series of repeated measures ANOVAs revealed that, while both groups improved their pronunciation in stress production in the posttest, the improvement seemed more pronounced in the explicit group.

In addition, 12 native English speakers rated comprehensibility and accentedness of the participants' extemporaneous productions on a 7-point Likert scale. However, no significant improvement was witnessed in either of these two aspects from the pre- to the posttest in both groups. Finally, a satisfaction survey utilizing 5-point Likert scale questions collected participants' views toward the training course. The combination of quantitative and qualitative results provides credence for the feasibility of instructing suprasegmental aspects of speech in language classrooms and highlights the advantages of explicit instruction.

Resumen

Este estudio tiene como objetivo explorar el efecto de la enseñanza explícita e implícita para mejorar la pronunciación del inglés en lo que respecta a la producción del acento. 28 alumnos chinos de inglés como lengua extranjera a nivel terciario se dividieron por igual en dos grupos y asistieron a un curso de formación de pronunciación de ocho sesiones de 30 minutos, impartidas utilizando los dos enfoques en cuestión respectivamente. Un pretest y un postest, que incluían un ejercicio de lectura controlada y otro de habla espontánea, se realizaron antes y después del experimento.

Las grabaciones de los participantes fueron analizadas acústicamente en PRAAT. Se obtuvieron lecturas de picos de intensidad, medias de F0 y duración para cada sílaba a nivel de palabra y para cada palabra a nivel de oración. Seguidamente se procesaron estos valores para obtener la diferencia en amplitud y F0, junto con las relaciones de duración, entre sílabas/palabras tónicas y átonas. Los resultados estadísticos generados a partir de una serie de ANOVAs de medidas repetidas revelaron que, si bien ambos grupos mejoraron su pronunciación en la producción del acento en el postest, la mejora parecía más pronunciada en el grupo explícito.

Además, 12 hablantes de inglés nativos evaluaron la comprensibilidad y el acento extranjero de las producciones extemporáneas de los participantes en una escala Likert de 7 puntos. Los resultados no revelaron mejora significativa en ninguno de estos dos aspectos en ambos grupos. Finalmente, una encuesta de satisfacción consistente en preguntas de escala Likert de 5 puntos recopiló las opiniones de los participantes con relación al curso de formación. La combinación de resultados cuantitativos y cualitativos proporciona credibilidad a la viabilidad de enseñar aspectos suprasegmentales del habla en las aulas de idiomas y destaca las ventajas de la enseñanza explícita.

x

Resum

Aquest estudi té com a objectiu explorar l'efecte de l'ensenyament explícit i implícit sobre la millora de la pronunciació anglesa pel que fa a la producció de l'accentuació. 28 estudiants xinesos d'anglès com a llengua estrangera a nivell terciari es van dividir de manera aleatòria en dos grups i van assistir a un curs de formació de pronunciació de vuit sessions de 30 minuts, impartida per utilitzar els dos enfocaments en qüestió respectivament. El pretest i el posttest, que implicaven una tasca de lectura controlada i una tasca de parla espontània, es van realitzar abans i després de l'experiment.

Els enregistraments dels participants es van analitzar acústicament amb PRAAT. Es van obtenir lectures de pics d'intensitat, mitjanes d'F0 i durada per a cada síl·laba a nivell de paraula i per a cada paraula a nivell d'oració. A continuació, es van processar aquests valors per obtenir la diferència d'amplitud i F0, juntament amb les relacions de durada, entre síl·labes/paraules tòniques i àtones . Els resultats estadístics generats a partir d'una sèrie d'ANOVAs amb mesures repetides van revelar que, mentre que els dos grups van millorar la seva pronunciació en la producció de l'accentuació en el posttest, la millora semblava més pronunciada en el grup explícit.

A més, 12 parlants natius anglesos van valorar la comprensió i el grau d'accent estranger de les produccions extemporànies dels participants en base a una escala tipus Likert de 7 punts. Els resultats no evidenciaren cap millora significativa en cap d'aquests dos aspectes dels dos grups. Finalment, una enquesta de satisfacció que utilitzava preguntes en base a escales tipus Likert de 5 punts va recollir les opinions dels participants en el curs de formació. La combinació de resultats quantitatius i qualitatius proporciona credibilitat per a la viabilitat d'ensenyar aspectes suprasegmentals de la parla a les aules d'idiomes i destaca els avantatges de l'ensenyament explícit.

Table of Content

List of Tables

List of Figures

CHAPTER I

INTRODUCTION

The role of pronunciation in enhancing the oral skills of L2 learners cannot be overstated. Achieving accurate and comprehensible pronunciation is essential for effective communication (Levis, 2015). However, the current state of teaching in English as a second language (ESL) and English as a foreign language (EFL) classrooms demonstrates that pronunciation instruction (PI) still falls short of desired standards. Despite an impressive wealth of empirical research demonstrating noticeable benefits of pronunciation instruction for EFL learners, many language teachers continue to perceive it as a major challenge. Surveys conducted among language teachers, including those in native English-speaking countries, have revealed a sense of trepidation and reluctance among L2 instructors when it comes to teaching pronunciation (see Foote et al., 2011). Particularly, there is a disparity among teachers regarding their priorities in the classroom, with some emphasizing grammar instruction while others place greater importance on vocabulary teaching.

Despite a substantial body of empirical research demonstrating the positive impact of PI on learners' pronunciation skills, a longstanding debate persists regarding the prioritization of phonological features, namely, segmentals versus suprasegmentals. Some proponents argue that suprasegmental features should take precedence in PI, as they exert a greater influence on speakers' overall comprehensibility and intelligibility (Field, 2005; Hahn, 2004; Zielinski, 2008). However, conflicting perspectives exist, which consider that difficulties with segmental features are often cited as a primary cause of communication breakdown (e.g. Collins & Mees, 2013). Zielinski (2015, p. 409) objectively asserted that segmental and suprasegmental features are both an integral part of the pronunciation system that can positively interact with and build on each other, hence they should not be viewed separately for the objective of improving English pronunciation as a whole. From a pedagogical standpoint, segmentals have received more extensive attention in language classrooms, primarily due to the

consistent perception that English suprasegmental features pose significant challenges in pronunciation learning.

In spite of calls from researchers and language practitioners to incorporate pronunciation instruction in the classroom, particularly focusing on prosodic elements, the responses of L2 learners indicate a lack of instruction or limited benefit from it (Derwing & Rossiter, 2002). Consequently, many teachers exhibit resistance to dedicating class time to this aspect. Furthermore, there remains a scarcity of reliable and comprehensive materials that specifically address prosodic aspects. Early research predominantly concentrated on segmental instruction, until in the recent decade a paradigm shift has been witnessed towards an increased emphasis on suprasegmentals and prosody. However, most textbooks and syllabi involving pronunciation teaching inadequately cover suprasegmental features unless specialized manuals are utilized (Busà, 2012, p. 101).

Furthermore, the effectiveness of explicit versus implicit approaches to teaching pronunciation in ESL/EFL contexts remains a subject of ongoing debate. The explicit approach involves exposing learners to explicit rules and utilizing their analytical skills to understand the mechanics of spoken language. In contrast, the implicit approach, rooted in behaviorist theories, emphasizes repetition and mimicry, aiming to encourage learners to intuitively grasp underlying rules without conscious awareness (Ellis, 2009). Studies conducted within pronunciation-teaching contexts have contributed to the existing controversy, with some favoring implicit instruction (e.g., Minhong & Ailun, 2006; Papachristou, 2011; Yang, 2016), while a larger body of researchers support explicit instruction. They conclude that explicit instruction proves more beneficial for L2 learners in improving their perception and production of L2 phonological features (e.g., Couper, 2003; Khanbeiki & Abdolmanafi-Rokni, 2015; Naeini & Adni, 2017; Yakut, 2020). Despite investigations into both approaches in pronunciation instruction, it is regrettable that a majority of studies have predominantly concentrated on the segmental level, with limited comprehensive exploration of prosody instruction. This deficiency in the literature serves as the driving rationale for this study.

1.1. Aims of the research

The current research intends to investigate whether Chinese EFL learners' pronunciation can be improved by a period of training as far as word stress and sentence stress are concerned. Explicit vs. implicit instruction methods are investigated in terms of their efficacy in improving EFL learners' pronunciation. Both approaches have been widely researched, and positive effects have been observed. This study aims to further explore whether improvements vary between the two approaches after the same period of training. The following two major research questions are formulated:

RQ1. Does prosody training help Chinese EFL learners with their production of English stress?

RQ2. Which approach (explicit vs. implicit) is more effective in improving Chinese EFL learners' production of English stress?

Additionally, it is also worth probing whether the training could impact learners' overall comprehensibility and accentedness as perceived by native speakers. Therefore, RQ3 is formed as follows:

RQ3. Do learners' perceived speech comprehensibility and accentedness show improvement following the training?

Grounded on the abundant previous research that indicates the efficacy of pronunciation training, the following hypotheses are made:

H1. Chinese EFL learners' pronunciation will improve in terms of word stress and sentence stress production after the experimental prosody training.

H2. Based on existing evidence, the explicit approach might work more efficiently than the implicit approach in a short-term training for improving EFL learners' production of stress.

H3. Learners' perceived overall comprehensibility and accentedness will improve, especially in the explicit group.

1.2. Structure of the dissertation

The dissertation comprises five chapters. After introducing the main theme and outlining the research goals in this chapter, Chapter Two provides a thorough theoretical background and literature review, divided into eight sections. Section 2.1 starts with a concise historical overview of pronunciation teaching methodologies. Section 2.2 delves into the differentiation between segmental and suprasegmental aspects of pronunciation and highlights the shift in teaching focus from the former to the latter. Section 2.3 subsequently provides an in-depth literature review of the successful implementation of suprasegmental instruction. Section 2.4 delineates the major components of suprasegmental features, ranging from word stress to different types of sentence stress. Section 2.5 embarks on a comparative analysis of Chinese and English rhythm, further addressing the contentious debate surrounding the existence of stress patterns in Chinese. Section 2.6 discusses the hierarchy of four pivotal acoustic cues for stress perception and production—intensity, pitch/F0, duration, and formant structure. Section 2.7 expounds upon the influence of native language (L1) transfer on second language (L2) pronunciation, supplemented by illustrative examples of challenges encountered by Chinese learners of English as gleaned from prior research. Concluding the chapter, Section 2.8 compares explicit and implicit instructions in detail, delving into an extensive literature review to evaluate the effectiveness of both methods in the context of pronunciation teaching.

Chapter Three presents the methodology employed for the experiment, detailing the research design, formation of participants, the materials utilized for both pretest and posttests, questionnaires, and, most importantly, the training materials created for the two experimental groups with the varied instruction. Afterwards, the data collection procedures and analysis methods are presented.

Chapter Four presents the results obtained from i) the acoustic analyses, ii) the raters' evaluations, and iii) insights garnered from the qualitative questionnaires. Starting with the quantitative results, acoustic measurements of intensity, F0, and duration are introduced in order, from analyses of word stress to sentence stress

respectively. Then, scores for comprehensibility and accentedness improvement are reported. Finally, students' responses to the Likert-scale questions are extracted and analyzed from the feedback questionnaires, followed by qualitative results from participants' responses to some open questions in each experimental group.

Chapter Five discusses the obtained results in relation to previous literature and finally concludes the whole study and reflects on the limitations encountered in the research. Simultaneously, the chapter elaborates on some pedagogical implications arising from the findings and suggests potential lines for further research.

CHAPTER II

THEORETICAL BACKGROUND

2.1. History of pronunciation teaching

The integration of pronunciation teaching into language classrooms could be traced back to almost two centuries ago. Murphy and Baker (2015) outlined four waves of pronunciation teaching, documenting its progress from the 1850s on. The first wave of language teaching prioritized the imitative-intuitive method, which dominated for more than three decades. This method emerged in stark contrast to the preceding grammar translation method, which focused primarily on reading and translating ancient literary works from languages such as Latin and Greek, with little attention given to pronunciation (Howatt & Widdowson, 2004; Kelly, 1969). Based on the belief that the most effective way to learn a new language involved a combination of imitation and intuition, the imitative-intuitive method relied on learners' ability to listen and reproduce the rhythm and sounds of the language without explicit instruction. Consequently, oral interaction between teachers and students became more prominent, leading to a greater recognition of the need to teach pronunciation. However, due to the lack of initial awareness regarding effective pronunciation instruction, classroom practice predominantly followed the principles of the imitative-intuitive method.

The second wave, spanning from the 1880s to the early 1900s, marked the emergence of the Reform Movement. This period saw the establishment of the International Phonetic Association (IPA) and the introduction of the analytic-linguistic approach. In contrast to the imitative-intuitive method, the analytic-linguistic approach emphasizes explicit instruction to achieve accurate perception and production of the target language (Celce-Murcia et al., 1996). Although originally intended for explicit teaching of grammar and language structures, the analytic-linguistic approach can be effectively adapted to include pronunciation instruction. This adaptation involves utilizing various resources, such as the phonetic alphabet, articulatory descriptions, vocal apparatus charts, contrastive information, and other aids, to enhance the teaching

of pronunciation alongside listening, imitation, and production skills. By incorporating these elements, learners can develop a more comprehensive understanding of the target language's phonological aspects and improve their accuracy in perceiving and producing sounds. Therefore, Celce-Murcia et al. (2010, p. 2) viewed this approach as a complement rather than a replacement for the intuitive-imitative approach.

Throughout much of the 20th century, these first two waves captured researchers' and language practitioners' attention, until in the mid-1980s, a third wave emerged with the rise and subsequent dominance of the communicative approach in language classrooms. The communicative approach placed a strong emphasis on interaction and real-life language use. It introduced authentic texts and situations into the learning process, aiming to achieve communicative goals. Unfortunately, this paradigm shift resulted in a neglect of pronunciation teaching and error correction. Various suggestions emerged during this time, indicating that errors could be overlooked during communicative activities (Brown & Yule, 1983) or that they should not be corrected at certain proficiency levels (Krashen & Terrell, 1983; cited in Levis & Sonsaat, 2017, p. 269).

These notions seem to suggest that pronunciation teaching was incompatible with the communicative approach. On the other hand, the main challenge was that very few teachers knew how to effectively integrate pronunciation teaching into communicative practices (Murphy & Baker, 2015). Some researchers explored the possibility of teaching pronunciation communicatively, but there still remained a lingering uncertainty about what exactly communicative pronunciation teaching should entail (e.g., Kenworthy, 1987; Morley, 1987; Pica, 1996). Just as Celce-Murcia (1983) suggested, teaching pronunciation communicatively is stepping into uncharted waters, particularly concerning suprasegmentals.

Despite considerable growth in the publication of textbooks focused on pronunciation during the 1980s and 1990s, particularly when compared with other areas of language teaching, there still remained a dearth of empirical research to substantiate researchers' methods and strategies. According to Murphy & Baker (2015), it was not

until the mid-1990s that empirical researchers began to explore foundational research questions aimed at supporting pronunciation instruction marking the advent of the fourth wave of pronunciation teaching. Among the initial studies conducted during this period was the research by MacDonald et al. (1994), who sought to compare the effectiveness of three different types of pronunciation instruction. Although the results were disappointing (as expected, considering the limited time frame), this study ignited significant interest in the empirical examination of classroom-based pronunciation instruction (see a review in Derwing, 2018).

2.2. Teaching pronunciation: from segmentals to suprasegmentals

Generally, English phonological features can be broadly classified into two categories: segmental and suprasegmental. Segmental features pertain to distinct units that can be physically or auditorily identified within speech, e.g., vowels and consonants. On the other hand, suprasegmental features are characteristics of speech that span across more than an individual sound in an utterance. Examples of such features are stress, rhythm, and intonation. These suprasegmental features utilize parameters like loudness, pitch, juncture, and duration (Sharma, 2021).

Traditionally, pronunciation materials or curricula primarily focused on small segmental elements. This pedagogical approach typically entails employing phonetic drills and repetition exercises to cultivate students' proficiency in accurately identifying, producing, and discerning these segmental sounds. By providing students with a solid foundation in segmental phonology, they are expected to be able to analyze and reproduce the intricate sound patterns of language, with the ultimate objective to improve pronunciation and phonetic accuracy in both speech perception and production. However, criticism has been directed towards the limited focus on segmental elements in pronunciation instruction, as the disregard for suprasegmental features can impede comprehension of how these elements coalesce within longer utterances (Pennington & Rogerson-Revell, 2019).

Suprasegmentals are widely acknowledged as the essential component in distinguishing semantic or pragmatic meanings and conveying speakers' emotions. A prime example lies in the placement of stress within a word, which can fundamentally alter its meaning and convey distinct concepts. For instance, the word "present" can denote a noun referring to *gift* when stress is placed on the first syllable, while, as a verb connoting *to report* when stress is shifted to the second syllable.

Additionally, introducing stress and rhythm within L2 classrooms can be beneficial to enhance students' communicative skills (Frost & Picavet, 2014; Kebboua Chaker, 2019; Quesada Vázquez, 2019; Tuan & An, 2010). For instance, sentence stress can reveal pragmatic meaning and emotions (a comprehensive overview of English stress is presented in Section 2.4 below). Stress placement is a crucial component in comprehension of L2 English speech. Failure to use rhythm properly can even cause communication breakdown (Jenkins, 2000). Consequently, proficiency in suprasegmental features encompasses not only the attainment of accurate pronunciation but also a deeper comprehension of a speaker's intended meanings. Many recent investigations into SLA have highlighted the critical importance of suprasegmental aspects in achieving fluency and proficiency in L2 interpretation and pronunciation, suggesting that learners who have greater exposure to prosodic aspects of language progress to spontaneous production better than those who have learned only segmental content (cited in Yang, 2016, p. 68).

2.3. Feasibility of suprasegmental instruction

Apart from the increasing awareness towards the importance of suprasegmental features, a great body of research has simultaneously witnessed positive effects of suprasegmental instruction. For instance, Derwing et al. (1997) investigated how explicit pronunciation instruction could help fossilized adult learners improve their intelligibility, comprehensibility, and accent in their speech production. 13 adult learners of English from diverse L1 backgrounds joined a speaking course twice per week over a period of 12 weeks. Due to the participants' varied native languages, the

10

focus of the course was primarily on suprasegmental features, as it was difficult to identify their common segmental problems. Before and after the instruction, participants read aloud a list of true and false statements, and their recorded utterances were evaluated by 37 native speakers through a transcription task to measure intelligibility. General comprehensibility and accent were assessed using 9-point Likert scales. Results indicated that all the utterances recorded in the post-instruction stage were more intelligible than those produced in the pre-stage. Interestingly, only the true sentences were rated significantly less accented and more comprehensible after the treatment. False sentences surprisingly did not show signs of improvement in terms of the studied constructs. The authors suggested that this discrepancy could be attributed to the greater unpredictability of false statements.

Subsequently, Derwing et al. (1998) expanded the research sample and conducted a ten-week study involving 48 adult L2 learners of English. The participants were divided equally into three groups of 16 students each. Different types of pronunciation instruction were implemented in each group: one experimental group focused on segmental instruction, another on suprasegmentals, while the third group served as the control group without explicit pronunciation instruction. Pre- and posttests were administered in the form of sentence reading and narration exercises with a picture prompt. The participants' readings were recorded and evaluated by native English listeners and trained ESL teachers. For the sentence-reading part, 48 native raters assessed participants' comprehensibility and accentedness using a 9-point Likert scale. Results indicated improvements in both segmental and suprasegmental instruction groups, while the control group showed no significant changes. Regarding the picture-prompted narration, 45-second excerpts were extracted and assessed by 6 ESL teachers in terms of comprehensibility, fluency, and accentedness. The suprasegmental group was the only group that showed improvement in all of the criteria. As a conclusion, Derwing et al. (1998) emphasized the importance of suprasegmental instruction, especially in spontaneous speech contexts, while acknowledging the positive role of segmental instruction in controlled situations.

In more recent studies, Tsiartsioni (2011) investigated the effectiveness of rhythm instruction in three different age groups of Greek learners (10, 13, and 16 years old respectively). For each age set, the experimental group received 25 pronunciation lessons with the focus on rhythm instruction. Vocalic and consonantal variability in a long stretch of speech was measured in the pre and post stage. Signs of a more native-like rhythm were shown in all age groups, while the control group remained stable.

Saito and Saito (2017) examined the effects of suprasegmental-based instruction on the global (comprehensibility) and suprasegmental (word stress, rhythm, and intonation) development of Japanese EFL learners. 10 students in the experimental group received suprasegmental-based instruction three hours per week over six weeks in total, while those in the control group were provided with meaning-oriented instruction without any focus on suprasegmentals. Speech samples elicited from read-aloud tasks were assessed by native-speaking listeners and acoustic analyses. After comparing the pre and post-test data, the experimental group showed significant gains in the overall comprehensibility, word stress, rhythm, and intonation in both trained and untrained lexical contexts.

Kebboua Chaker (2019) investigated the potential of using technology in the teaching and learning of English pronunciation with regard to word and sentence stress. Two groups of 24 first-year university students participated in the study—the control group received instruction on English stress by conventional means whereas the experimental group made use of an online tool for learning pronunciation. Following a pretest, intervention, and post-test design, the acoustic results of the study revealed that, while both the control and experimental groups showed improvement in their pronunciation of word and sentence stress, this improvement was more significant in the experimental group, even though there was a great deal of variability depending on the specific word or phrase. The results therefore provided evidence of the usefulness of information technologies in the classroom for teaching English suprasegmental features.

Li et al. (2023) explored the effects of embodied prosodic training on the production of non-native French front rounded vowels and the overall pronunciation proficiency. 57 Catalan learners of French practiced pronunciation in one of two conditions: one group observed hand gestures embodying prosodic features of the sentences they were listening to, while the other group did not see any such gestures. The learning outcome was assessed in a pretest, posttest, and delayed posttest through a dialogue-reading task and a sentence imitation task in terms of accentedness, comprehensibility and fluency scores, and through formant analysis of participant-produced target vowels. The results showed that using hand gestures to mimic prosodic features could not only enhance learners' overall pronunciation proficiency but also improve the accuracy of French front rounded vowels. Such a finding corroborates the hypothesis that prosodic and segmental features are interdependent in phonological learning and that instruction of prosodic features can act as a scaffold for improving articulatory production.

The success achieved in the aforementioned suprasegmental treatments highlights the feasibility and benefits of integrating suprasegmental teaching into language classrooms. It is possible to enhance learners' overall pronunciation through appropriate training techniques. The extensive body of literature provides a robust empirical foundation for implementing the interventions targeting English stress carried out in the current study.

2.4. What suprasegmental features to teach

Bearing in mind the intricacy of suprasegmental content, it is not realistic to cover all aspects within most standard English language classrooms. Thus, in the current study, stress, including word stress and sentence stress, is the primary focus. The subsequent sections will elaborate on the distinct patterns of word stress and sentence stress, respectively, with particular attention to English.

2.4.1. Word stress

According to Trask (1996, p. 336), stress is the relative emphasis or prominence given to a certain syllable in a word or to a certain word in a phrase or sentence. It is the first and foremost component of suprasegmental features, along with pitch accent, tone and intonation. Liberman (1975) proposed that syllables are perceived as stressed based on their relationship with non-stressed syllables. In essence, stress is a relative characteristic that emerges from the interplay between two or more syllables, rather than an inherent property of a single syllable itself.

A stressed syllable is typically marked by increased loudness, vowel length, full articulation of the vowel, and changes in tone (Fry, 1955). Laver (1994) further characterized stress as follows: "Other things being equal, one syllable is more prominent than another to the extent that its constituent segments displays higher pitch, greater loudness, longer duration or greater articulatory excursion from the neutral disposition of the vocal tract" (p. 511).

According to Mott (1991, p. 215-216), word stress in the world's languages can be broadly categorized into two types: fixed-stress languages and dynamic-stress languages. Some languages have fixed stress, meaning that the stress always falls on the same position within a word. For example, among these languages, Czech and Finnish words always carry stress on the first syllable, while Turkish, Persian, and modern Hebrew have the final syllables stressed. In dynamic-stress languages like English, Spanish, and Russian, the position of stress within a word is not fixed or predictable and it is usually lexically encoded. In many cases, various levels of stress, such as *primary stress* and *secondary stress*, can be identified. However, it is important to note that stress is not considered a universal feature of languages. Hyman (1977) examined 444 languages and suggested that about 2/3 of them have stress and the rest have no stress. Particularly, controversial arguments exist regarding the presence of stress in Mandarin Chinese (Duanmu, 2000). The major argument is that Chinese is a tonal language rather than a stress-based language and it relies on the pitch contour or tone of a syllable, instead of lexical stress, to differentiate semantic meaning. Thus,

Chinese phonological features are considered to be significantly different from English (see further details of Chinese and English rhythms in comparison in Section 2.5).

In English, lexical stress is typically categorized into three levels of prominence: primary stress, secondary stress, and unstressed syllables. Primary stress represents the strongest syllable in a word, and it is indicated in IPA by a raised vertical line [ˈ] at the beginning of the syllable, as seen in the word "student" /ˈstudnt/. Secondary stress is stronger prominence relative to unstressed syllables but not as much as a syllable with primary stress, marked with a lowered vertical line [ˌ]. Unstressed syllables are commonly realized with a schwa sound without any marking symbol. A clear illustration of the distinction between secondary stress and the complete absence of stress can be observed in the final syllable of "delegate", where the verb form is pronounced as /ˈdɛləˌget/ and the noun form as /ˈdɛləgət/.

Stress plays a crucial role in speech intelligibility and effective communication. Kenworthy (1987) regarded stress as an essential feature of word identity in English, as one major function of word stress is differentiating homographs, such as "**re**cord" as a noun and "re**cord**" as a verb[1]. Misplacing word stress may obfuscate the word and lead to misunderstanding. For instance, in a study on the processing of lexical stress, Cutler and Clifton (1984) found that incorrect stress placement in disyllabic words have detrimental effects in speech processing. A shift of stress from the left syllable to the right seriously hindered intelligibility.

Despite a lack of regulated stress patterns as in the fixed-stress languages, English word stress is not entirely capricious. There are some predictable patterns or paradigms that can help students anticipate stress placement. The following selected rules, adapted from Hewings (2007), can assist in predicting stress placement[2].

• Most two-syllable nouns and adjectives have stress on the first syllable, for example, "**fa**mous", "**re**cord", "**dri**ver". However, there are some exceptions to this

---

[1] Syllable in bold indicates the primary stress position.

[2] The rules of word stress placement are not the focus of the present research; thus, they were not taught in the experimental training course.

15

general rule, e.g. "ga**rage**", "ma**chine**", especially when prefixes are involved which are typically distressed, e.g. "a**sleep**", "mis**take**".

• On the other hand, two-syllable verbs in English, by and large, possess the primary stress on the second syllable, e.g. "re**cord**", "pre**sent**". Likewise, exceptions exist too, such as "**pro**fit", "**en**ter", "**che**rish".

• Compound nouns that combine two smaller words usually exhibit stress on the first word, e.g. "**book**shop", "**bus** stop", "**tra**ffic light", whereas most compound adjectives, when both parts are stressed, show stronger or primary stress on the second word, e.g. "well-**done**", "good-**looking**". Besides, the majority of compound single-word verbs also share primary stress on the second syllable, e.g. "over**hear**", "double**check**".

• However, when compound words are incorporated to phrases, stress may switch to other positions, for example, "a **good**-looking boy", "a **well**-done project"[3].

2.4.2. Sentence stress

Sentence stress, also known as *prosodic stress*, refers to stress patterns that apply at a higher level than individual words, namely, within a prosodic unit. It involves the placement of varying degrees of emphasis or prominence on specific words within a sentence, highlighting their significance and aiding in conveying the intended meaning (Clark & Yallop, 1995). Sentence stress, in fact, is the feature that creates the rhythm of English. It is achieved through the combination of stressed, lightly stressed, and unstressed syllables, which creates the distinctive musicality of English. Particularly, for the rhythm of English to sound natural, stressed syllables within a sentence should stand out whereas unstressed syllables should be "squeezed" in between the stressed ones. This consistent timing and interplay between stressed and unstressed syllables generate the rhythmic beat of English. Section 2.5 will elaborate on this stress-timed English rhythm in comparison with the syllable-stressed rhythm of Chinese. Prior to

---

[3] More stress rules on the sentence level can be found in the subsequent section.

the comparison, some crucial concepts and elements related to English sentence stress will be systematically presented in the following order.

2.4.2.1. Content words vs. function words

In an English sentence, content words, including nouns, adjectives, main verbs, adverbs etc. are typically stressed by default, as they bear semantic meaning and convey important information within a sentence. On the other hand, function words, such as auxiliaries, articles, prepositions, pronouns, and conjunctions, are less likely to be stressed, as they serve as grammatical elements that connect and structure a sentence rather than carrying independent semantic content (Prator & Robinett, 1985). Table 1 provides the types of content words and function words along with examples in English.

Table 1

List of content words and function words in English

Types of content words	Examples	Types of function words	Examples
Nouns	boss, family	**Auxiliary verbs**	do, have
Main verbs	run, come	**Conjunctions**	and, or
Adjectives	beautiful, smart	**Prepositions**	for, on
Adverbs	carefully, slowly	**Articles**	a, an, the
Negation words	not, never	**Pronouns**	he, she
Interrogative	when, who	**Relative pronouns**	who, that
Demonstratives	this, that	**Possessive adjectives**	his, her
Numerals and ordinals	one, second	**One used as a noun-substitute**	the blue one

Note. Modified from Prator & Robinett, 1985.

However, it is important to note that not all content words carry the same significance with respect to the meaning of a syntactic structure. Consequently, the

prominence or weight assigned to different content words may vary. In other words, the hierarchy of sentence stress is determined by the relative importance of individual words within the sentence. Despite the fact that sentence stress itself does not alter the meaning of any specific lexical items, it has been widely acknowledged to play an essential role in facilitating intelligible communication in English (Jenkins, 2002).

2.4.2.2. Tonic stress placement

Generally, English sentence stress is classified into two types of tonic stress placement: *neutral tonic placement* and *dislocated tonic* (Ladd, 1996). The former, also known as *unmarked sentence stress*, does not express emphasis or contrast. In particular, when all the information conveyed in a sentence is new, the neutral tonic stress occurs towards the end of a phrase/sentence by default, falling on the last content word, as shown in example (1) below. The syllables in bold carry the tonic stress within the tonic words. This type of stress is found out to exist in up to 90% of spoken phrases/sentences in English (Crystal, 1969).

(1). The computer is **use**ful.

Nevertheless, in many circumstances, the tonic syllable in a sentence can be dislocated and shifted to the nearest preceding content word or even function word. Bardovi-Harlig (1986) summarized that the focus on new information, contrast and emphasis serves as the driving force behind the shift of sentence stress. For marking new information, neutral tonic stress may be moved frontwards when the last content word conveys old information, as exemplified in (2).

(2). Person 1: Shall we have dinner this **Fri**day?

Person 2: Sorry, I don't have **time** this Friday!

In terms of contrastive stress, it is employed to call attention to contrasts. A contrast may be explicitly presented, as in (3); or implied implicitly as in (4).

(3). I want the **blue** car, not the **ye**llow car.

18

(4). I want the **white** phone.

Lastly, emphatic stress is used to accentuate a certain word in a sentence. Placement of emphatic stress is unpredictable, as its position largely depends on a speaker's intention or emotion (Bolinger, 1972). As shown in (5), stress shifts from the final content word "tried" to "best" to highlight the speaker's feeling.

(5). The food is the **best** I've tried.

Sometimes, even a function word can be stressed as long as it conveys weighty meaning (Kuiper & Allen, 1996), as in (6).

(6). Person 1: You weren't at the meeting?

Person 2: I **was** there (auxiliary word as the emphasis)

2.5．English and Chinese rhythm in comparison

English and Chinese represent two distinct rhythmic patterns: stress-timed versus syllable-timed. According to Abercrombie (1967), stress-timed languages, including English, exhibit a rhythm where the time intervals between stressed syllables tend to be relatively equal. Stressed syllables are emphasized and tend to be pronounced with longer durations, while unstressed syllables are reduced or shortened. This rhythmic pattern creates a sense of regularity and predictability based on the timing of stressed syllables. By contrast, in syllable-stressed languages, such as Chinese, each syllable is given roughly equal duration or time, regardless of its stress status (Mackay,1985). Figure 1 illustrates this contrast between stressed-timed English rhythm and syllable-timed Chinese rhythm. In English, stressed syllables have relatively equal timing intervals between them, creating a rhythmic pattern. On the other hand, in the Chinese sentence, each word is pronounced with the same degree of stress and the same amount of time. This results in a more uniform timing of syllables, without the distinct timing variations between stressed and unstressed syllables found in stress-timed languages.

Figure 1

Examples of English and Chinese sentence rhythm
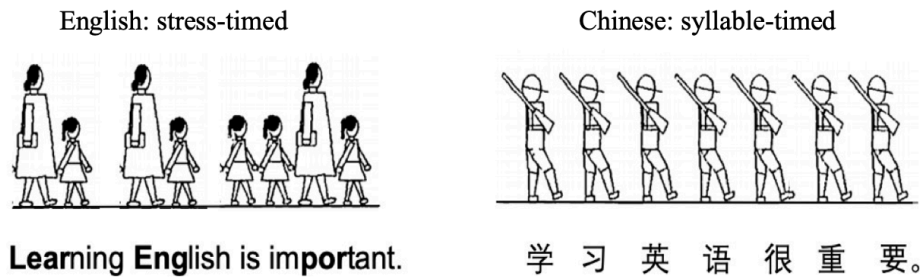
Note. This figure is modified from Prator, & Robinett, 1985

Wang et al. (2003) conducted further research on the disparity between the two languages in terms of stress patterns. According to their findings, words in an English sentence are typically considered unstressed by default, with stress serving to mark important meanings and enhance semantic or pragmatic functions. In contrast, in Mandarin, words in a sentence are generally viewed as stressed by default, and the linguistic functions mainly depend on de-stressing instead of stressing. Similarly, Třísková (2016, p. 100) suggested that the non-stress assignment should be investigated rather than stress assignment in Mandarin Chinese. Hu (2017) summarized that "Chinese rhythm represents itself in the collocation among its various tones, while English rhythm is realized by the alternative appearance of the stressed and unstressed syllables" (p. 408).

However, while there is extensive research on stress in English, it remains one of the most elusive areas in Chinese Phonology. The presence of stress in Chinese is still a topic of ongoing research and controversy exists among linguists regarding its existence. Some linguists have provided subtle judgments on stress in Mandarin. For instance, Chao (1968, p. 35) suggested that, in a string of full syllables, the last syllable carries the most stress, the first one has slightly less stress, and the rest of syllables have even less. This is known as the *medium-light-heavy* theory. However, Chao (1968, p. 38) also acknowledged, Mandarin speakers do not agree on where stress is, making it difficult to confirm the medium-light-heavy theory. While stress is not considered a

20

primary feature of Chinese, some scholars argue that there are elements of stress-like phenomena present in certain contexts. The following sections will introduce the Chinese tone system with examples and further discuss some stress-related phenomena in Chinese.

2.5.1. Chinese tones

As has been shown, languages differ not only in their inventory of segmental contrasts but also in their suprasegmental properties. Chinese is distinguished notably from English as a tonal and syllable-stressed language. As remarked by Xu (2019), "The key to the difference (between English and Chinese phonology) is that polysyllabic words are the domain of English stress assignment and monosyllabic words are the domain of Chinese tone assignment" (p. 54). Modern standard Mandarin, specifically, has five distinct linguistic tones, each indicated by diacritic marks above the vowels in pinyin — high level (Tone 1, -), rising (Tone 2,´), falling-rising (Tone 3,ˇ), falling (Tone 4, `), and neutral tone (no specific contour). A change in the tone of a word can give rise to different lexical meanings. For example, the syllable 'ma' can be pronounced in any of the five tones: mā, má, mǎ, mà, ma, each corresponding to different Chinese characters such as 妈 (mum), 麻 (numb), 马 (horse), 骂 (scold) and 吗 (an interrogative particle). Variations in pitch level primarily contribute to the prominence of a word or sentence in Mandarin (Bian, 2013, p. 203).

Additionally, monosyllabic or disyllabic words form a significant portion of the Chinese language (Chen, 1984). Each morpheme is independent and meaningful, carrying distinctive inherent pitch contours. Due to this characteristic, Chinese is generally considered to lack a strong pattern of stress, both at the segmental and suprasegmental levels. However, it is worth noting that there are certain contexts where elements resembling stress can be observed in Chinese, particularly when words are combined to form larger units or sentences. Exploring the assignment of stress in Mandarin is a complex but valuable area of study, as understanding both stressing and de-stressing aspects contributes to a comprehensive understanding of the language's

rhythm and pronunciation patterns. The following section will provide examples of these contexts that illustrate stress patterns in Chinese.

2.5.2. Evidence of stress in Chinese

Chao (1968, p. 35) identified three types of stress patterns in the phonological system of Mandarin Chinese: *contrastive stress*, *normal stress*, and *weak stress*. Contrastive stress can be seen in the cases of the Mandarin words *chaofan* vs. **chaofan** (see Wang, 2008). With the same segmental contents and exactly the same tones on both syllables, the first item carries the stress on the second syllable and represents "stir-fried **rice**" (but not noodles or vegetables), while the second item with stress on the first syllable means "**stir-fried** rice" (emphasize the cooking method). An exaggerated pitch range and greater intensity can be observed on the stressed syllable, on top of the innate tone pitch of each morpheme. The normal stress and weak stress mean no exaggerated pitch or intensity is used. The sentence example in (7) illustrates the positions of contrastive stress, normal stress, and weak stress. Contrastive stress is on 书 (books), which is the tonic stress in the sentence too, to emphasize what is bought, 我 (I) and 买 (bought) carry the normal stress determined by their pitch, and the final *le* is a modal particle that has weak stress.

(7). *我买书了。*

I bought **books**.

With regard to stress assignment in Mandarin, it is proposed in many studies that the change of pitch intensity in Mandarin is usually placed in the final position of a sentence. For example, Hara et al. (2014) detected Mandarin speakers' tendency to use sentence-final stress to indicate the certainty in their propositions. Likewise, Wang and Chu (2006) investigated the distribution of stress in 300 freely produced Chinese sentences. The finding indicated that the stressed syllable tends to lie in the final position of a sentence, assigned to i) the predicate in a subject-predicate structure; ii) the object in a verb-object structure; iii) the head in an adjunct-head structure unless the

head is highly predictable. They are exemplified as below:

(8). i) subject-predicate structure

*两国之间的关系前景**美好**。*

The prospect of the relationship between the two countries is **good**.

ii) verb-object structure

*这种言论只会引起社会**混乱**。*

Such kind of speeches can only cause social **disorder**.

iii) adjunct-head structure

*国际清算银行将发挥**重要**作用。*

BIS will play an **important**[4] role.

He (2002) provided an in-depth contrastive analysis of sentence stress in English and Mandarin. He discovered that, similar to English, sentence stress in Mandarin is driven by grammatical, semantic and emotional factors and it can be classified into two categories—grammatical stress and emphatic stress. In Mandarin, the change in sentence stress serves numerous pragmatic functions that are comparable to the contrastive or emphatic stress observed in English. Some examples illustrating this are provided in (9) and (10).

(9). A: *你要这辆**黄色的**车吗？*

Would you like the **yellow** car?

B: *不，我要**绿色的**车。*

No, I want the **green** car.

(10). *他**真的**不是故意的。*

He **really** didn't do it on purpose.

---

[4] *作用* (role) functions as the head of the object noun phrase. Given its frequent occurrence as the object of the verb *发挥* (play), and usually modified by the adjective *重要* (important), it becomes highly predictable based on the two preceding words. Consequently, emphasis is directed towards *重要* rather than *作用*.

Nevertheless, it is worth noting that emphatic/contrastive stress in Mandarin could also be manifested by using "是" cleft (Chinese verb-*be*) as a stress indicator instead of utilizing extra pitch for certain contrastive words (Li, 2008). For instance, as shown in (11).

(11). *他喝的牛奶，不是果汁*

He drank milk, **not** juice.

In addition to the arguments presented above, there have been perceptual and acoustic studies that provide further evidence for the presence of sentence stress in Mandarin (e.g. Chao, 1968; Shen, 1993; Wang, 2008; Xu, 1999). Chao (1968) claimed that, similar to stress in English, pitch, intensity, and duration are the three factors that lead to the realization of stress in Chinese, although prominence in Chinese primarily relies on pitch variations. Shen (1993) conducted an acoustic study on sentence stress in Mandarin and reported that the stressed words produced by Mandarin speakers exhibited higher intensity while unstressed words had lower intensity. Wang (2008) further confirmed that stress in Chinese shares some characteristics with English, such as higher pitch, greater intensity, and longer duration, but noted that the relative importance of these cues differs between the two languages. Although these studies do not provide specific information about the exact positions of stressed syllables within a sentence, they collectively confirm the existence of sentence stress in Mandarin. The subsequent section will review acoustic studies pertaining to stress perception and production involving both L1 and L2 speakers, in a manner that considers the potential hierarchy of importance among the four well-established acoustic cues.

2.6. Acoustic correlates of stress

Traditionally, several auditory cues have been correlated with word stress, including: (i) loudness, (ii) length, (iii) pitch, and (iv) quality. As defined by Laver (1994), "Other things being equal, one syllable is more prominent than another to the extent that its constituent segments display higher pitch, greater loudness, longer

duration or greater articulatory excursion from the neutral disposition of the vocal tract" (p. 511). From the acoustic perspective, these perceptual parameters correspond to intensity or amplitude, duration, fundamental frequency (F0) and vowel quality, respectively (Terken & Hermes, 2000). These parameters play a crucial role in the perception of stress, and are manifested in the speech patterns of both native and non-native speakers. Nevertheless, research has demonstrated that native speakers of English execute a more intricate combination of F0, duration, and intensity to convey lexical stress. Conversely, L2 learners tend to transfer their native prosody into their English speech production (Guo, 2022).

Yet, abundant experimental work has demonstrated that these four acoustic factors do not have equal importance. It is widely believed that the primary cue for stress in English, both in natural speech and synthesized speech, is relative pitch prominence (F0). Beckman (1986) proposed that F0 contour outranks amplitude contour (intensity), duration, and spectral quality (vowel quality). Roach (2009), in the same vein, stated that the most significant effect is produced by pitch, followed by length and amplitude.

Besides, F0 is also considered the primary acoustic cue for Mandarin tones (e.g. Altmann, 2006; Fu & Zeng 2000; Lin, 1988). Lin (1988) conducted an experiment with synthesized speech, and he found that Mandarin speakers primarily rely on F0 height and F0 contour for tone perception. Any variation in duration or intensity, under the circumstance of a constant F0, does not affect the tone perception. Besides, Xu (1999) also found Mandarin speakers vary F0 during the production of Mandarin phrases to differentiate stressed and unstressed words. According to Altmann (2006), "We do not observe the combination of pitch, duration and intensity referred to previously as the manifestation of stress (in Mandarin Chinese). Instead, pitch alone typically provides crucial word level contrasts" (p. 46).

However, the investigation of F0 as a defining factor for English lexical stress is notoriously challenging due to the co-occurrence of word-level prominence and post-lexical tonal events, such as pitch accents (Beckman, 1986; Ladd, 2008). This means that the pitch of stressed syllables may vary depending on whether the syllables carry a

phrase-level accent (Chrabaszcz et al., 2014). Moreover, phrase-level accent is determined by the communicative context, which influences the placement of the highest prominence within a phrasal unit or sentence (Eady & Cooper, 1986). For example, stressed syllables in accented words receive more prominence than stressed syllables in unaccented words. Grounded on that, it is argued that F0 is the trickiest and least reliable acoustic correlate in comparison to duration, vowel quality, or intensity in the perception of English lexical stress.

On another note, according to Gordon and Roettger's (2017) cross-linguistic survey of 110 (sub-) studies on acoustic correlates of word stress, duration is by far the most frequently measured property in the database across 72 languages and the most successful marker of stress, distinguishing stress in 85 of 100 (sub-) studies. Chun (1982) investigated the perception of sentence stress in Mandarin phrases and found that sentence stress could be realized through either differential use of F0 or lengthening of syllable duration, depending on the position of the stressed word within the sentence. In Jung and Rhee's (2018) cross-linguistic study of Korean, Japanese and Chinese speakers producing English lexical stress, the results demonstrated that non-native English speakers were able to produce English stress in a similar way to native English speakers, and all speakers (both native and non-native) showed a tendency to use duration as the strongest cue in producing stress.

It has long been believed that stressed vowels typically have higher intensity than unstressed vowels (Beckman, 1986; Fry, 1955; Shen, 1993), indicating that they are produced with more energy, usually with an open vocal tract and voicing. Shen (1993) confirmed that the stressed words produced by Mandarin speakers exhibit higher intensity while unstressed words have lower intensity. However, the role of intensity in predicting stress has been a subject of debate in the literature. Fry (1955, 1958) conducted a series of studies to explore the influence of the accoustic correlates in English stress perception by native speakers and found that vowel duration and intensity are the most correlated with perceived stress. Changes of vowel duration ratio

influenced the listeners' stress judgements. Despite a similar effect detected in intensity ratio, it did not cause a complete shift in stress judgement.

Finally, formant structure has also been found to be relevant for English stress perception and production. However, it is often considered as the weakest cue for stress perception compared to F0, duration and amplitude (Fry, 1965) and the least researched.

Overall, establishing the hierarchy of the four acoustic cues that contribute to stress is a complex issue. Stress is not determined by a single mechanism, but is a composite of all these factors, and its understanding requires consideration of all the aforementioned elements.

2.7. L1 interference in pronunciation

There has been a great deal of research investigating the impact of the native language on foreign language learning. According to Odlin (1989), "Transfer is the influence resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired" (p. 27). Specifically, where the two languages are identical or similar, positive transfer of the native language pattern takes place, making learning easier (Ringbom, 2007). Conversely, if L1 and L2 prosodic elements differ substantially, the chance for learners to produce L2 prosodic features incorrectly is higher, which contributes to a noticeable degree of foreign accent in production (Munro & Derwing, 1995; Trofimovitch & Baker, 2006).

Research has consistently demonstrated that L2 learners often struggle to accurately reproduce the distinctive features of L2 sounds and their allophonic variations. The negative transfer of L1 features can result in inaccuracies and errors in L2 production. When L2 learners attempt to articulate sounds that do not exist in their L1, they may inadvertently replace the target sounds with similar ones from their native language. For example, as observed by Busà (2008), Italians who are trying to differentiate between words like "bead" and "bid" may exaggerate the duration of the vowel in "bead" deliberately to create distinctions like /bi::d/ and /bid/. This

phenomenon highlights how L2 learners may resort to using phonetic features from their L1 as substitutes in order to bridge the gap between the unfamiliar L2 sounds and their existing linguistic knowledge.

Likewise, the phonological features in Chinese also exert an influence on Chinese learners' acquisition of English stress. It has been revealed in multiple studies that Chinese speakers encounter difficulties in acquiring English lexical stress because of their tonal L1 background. For instance, some Chinese learners tend to apply tones to indicate different English lexical meanings (Bian, 2013). Juffs (1990) found that Chinese speakers use Tone 1 with excessive length to indicate lexical stress. Cantonese speakers, on the other hand, tend to associate high and low tones with stressed and unstressed syllables (Chao, 1980). In Archibald's (1997) study, Chinese subjects appeared to treat stress as a purely lexical phenomenon, similar to tone, rather than understanding the principles of English stress assignment. These difficulties have been attributed to the transfer of prosodic features of the Chinese tonal system.

Moreover, the placement of sentence stress in production also poses a consistent challenge for Chinese EFL learners. Zhang et al. (2008) observed that native Mandarin speakers, due to the influence of syllable-timed rhythm in their native language, tend to preserve syllable-timed rhythm and fail to adequately reduce vowels in unstressed syllables. Hismanoglu (2009) also suggested that learners who have a syllable-timed mother tongue (e.g. Chinese) often overlook stress patterns or incorrectly allocate stress in English words or sentences, unless they are exposed to features of English stress allocation. Low-proficiency Chinese learners tend to place sentence stress on most of the words in a sentence, even on function words without realizing that their native language affects their English speech rhythm, leading to unnatural or incomprehensible utterances (Lee, 2013).

In terms of L1 transfer acoustically, numerous studies have explored how acoustic attributes of prosodic contrasts in the L1 can affect perception of prosody in the L2 (e.g. Best & Tyler, 2007; Chrabaszcz et al., 2014; Flege & MacKay, 2004). Chrabaszcz et al. (2014) summarized that such an impact is displayed in at least two different ways—

the stress pattern in a word (or the type of tone in tonal languages) and the acoustic cues used to realize prosodic contrasts. Evidence to date reveals that speakers are adept at using acoustic cues in the L2 if the same types of cues are actively used in the L1. For example, Zhang et al. (2008) also observed that native Chinese speakers use fundamental frequency as the dominant cue to signal stress in L2 English utterances, which could be attributed to the L1 influence. As previously mentioned, F0 is the most important acoustic parameter of Chinese tone (e.g. Altmann, 2006; Fu & Zeng, 2000; Lin, 1988).

Similarly, native speakers of Vietnamese can perceive stress contrasts in English despite the differences in prosodic structures between the two languages because Vietnamese phonology already includes important elements involved in English stress contrast, such as pitch cues and intensity (Nguyen et al., 2008). However, the same study reveals that beginner learners' production of other inactive elements, such as duration contrast and vowel reduction, are incomplete or absent in most cases, whereas the advanced speakers could produce native-like duration contrasts. Keyworth (2015) justified that native-like command is achievable through a large amount of L2 input and increased practice.

2.8. Explicit vs. implicit instruction

The issue of whether to teach pronunciation implicitly or explicitly has long been controversial. According to Ellis (2009), explicit instruction constitutes direct intervention and provides descriptions and explanations of the rules to facilitate learning. On the other hand, implicit instruction advocates unconscious and automatic learning, which is similar to the process of first language acquisition. It encourages learners to infer underlying rules without awareness (Ellis, 2009). As Norris and Ortega (2000) explained, "When neither rule presentation nor directions to attend to particular forms are part of a treatment, that treatment is considered implicit" (p. 437). Some typical implicit techniques of instruction include listening-only interventions and shadowing, where learners imitate a presented speech stimulus as closely and

quickly as possible (Derwing & Munro, 2014). Foote and Mcdonough (2017) held a favorable view of this approach, asserting that this instructional method provides learners with valuable opportunities to discern disparities between their own pronunciation and that of their target language. Consequently, this facilitates the mitigation of L1 influence and individual differences to a considerable extent. Table 2 provides an overview of the primary attributes of implicit and explicit language instruction.

Table 2

A comparison between implicit and explicit language instruction

Implicit Instruction	Explicit Instruction
Attracts attention to target form	Directs attention to target form
Is derived spontaneously	Is predetermined and planned
Is unobtrusive (minimal interruption of communication of meaning)	Is obtrusive (interruption of communication of meaning)
Presents target forms in context	Presents target forms in isolation
Makes no use of metalanguage	Uses metalinguistic terminology
Encourages free use of the target form	Controlled practice of the target form

Note. This table is based on Housen & Pierrard (2005), cited in Peltekov (2020).

While previous research clearly demonstrates that explicit grammar instruction is more effective than implicit grammar instruction (see Norris & Ortega, 2000; Spada & Tomita, 2010 for meta-analyses), the studies exploring the effects of implicit and explicit instruction on pronunciation are still far from reaching conclusive results. A review of the literature in the recent decades shows that more studies carried out in EFL/ESL contexts have taken side for the explicit approach and obtained positive results generated from explicit teaching (e.g. Ghorbani et al., 2016; Gordon et al., 2013; Khanbeiki & Abdolmanafi-Rokni, 2015; Naeini & Adni, 2017; Yakut, 2020). On the other hand, fewer studies have explored or favored the implicit instruction of

pronunciation (e.g. Minhong & Ailun, 2006; Papachristou, 2011; Shamiri & Farvardin, 2016).

2.8.1. Explicit instruction

Derwing and Munro (2005) focused on the effectiveness of explicit instruction on mutual intelligibility and comprehensibility. Their study found that explicit teaching of pronunciation improved mutual intelligibility and comprehensibility of the learners. They concluded that, when appropriate instruction was not given, a foreign language accent might reduce intelligibility in both native-nonnative and nonnative-nonnative interactions.

Similarly, Saito (2011) investigated the effects of explicit phonetic instruction on second language pronunciation by measuring comprehensibility and accentedness. Twenty native Japanese learners of English in ESL settings were randomly assigned to the experimental group and the control group. The participants in the experimental group received explicit instruction on target segmentals, while those in the control group did not receive direct pronunciation training. Results from four native English judges suggested that explicit instruction has a significant effect on comprehensibility especially in sentence-reading tasks, although a significant reduction in foreign accent was not detected in any contexts.

Explicit instruction on suprasegmentals has also been proven effective. Gordon et al. (2013) conducted a classroom-based study to investigate how explicit instruction influences the acquisition of L2 phonological features, and how their production contributes to comprehensible speech in L2 learners. Three groups of ESL learners received pronunciation instruction for three weeks. Two experimental groups received explicit instruction either on suprasegmental or segmental features. A combination of these features was presented orally to the third group without explicit instruction. The pre-test and post-test recordings indicated that only the explicit group trained on suprasegmentals improved its comprehensibility scores significantly from pretest to posttest. The authors advocated that explicit phonetic instruction can make learners

notice L2 features, particularly, through explicit presentation of contents, guided analysis and practice, corrective feedback, and so forth, so as to reach the goal of facilitating L2 learners' development of comprehensible speech.

Quesada Vázquez (2019) examined the efficacy of incorporating explicit rhythm instruction into the English language classroom for the improvement of students' comprehensibility and fluency. The study involved first-year engineering students enrolled in a technical English course, who engaged in 10 weekly pronunciation modules structured under a communicative framework. The participants were divided into two groups: an experimental group that received explicit rhythm instruction, and a control group that did not receive such instruction. Students were recorded before and after the treatment and their performances were compared to examine their progress. The acoustic analysis focused on VarcoV values and sentence pauses, while native English speakers evaluated the comprehensibility and fluency of the students' extemporaneous productions. The findings revealed differences between group means and significant effect sizes for the acoustic measurements, although not all general results reached statistical significance. The combination of quantitative and qualitative results suggest that providing explicit instruction in suprasegmental aspects of speech, such as rhythm, can be beneficial for language learners.

Among the studies which have concluded that explicit instruction outweighs implicit instruction, Chung (2008) compared explicit and implicit instructions for improving Chinese learners' production of English word stress and found that both groups improved equally in the posttest, but the explicit group showed more advantages in the delayed posttest. These results have also called for a stronger role of phonetic explicitness in L2 pronunciation instruction.

Khanbeiki and Abdolmanafi-Rokni (2015) investigated the efficacy of explicit vs. implicit instruction on the learnability of English consonant clusters. 60 intermediate EFL learners were evenly put into three groups. One of the groups received explicit instruction including exposure to the correct pronunciation of consonant clusters, extra examples, and explanations. The second group received implicit instruction, including

32

only exposure to the correct pronunciation of such clusters. The third group, serving as the control group, received no instruction on pronunciation. After ten sessions of instruction, all the participants were provided with the posttest in which they had to read aloud a series of words containing initial and final consonant clusters. The results showed that although they were not significantly different in the pretest, they performed differently on the posttest. Despite the visible progress from the implicit group during the course, explicit instruction was the best method in improving participants' pronunciation, specifically, explicit instruction was better than the implicit instruction, followed by the control group.

In the same vein, Yakut (2020) provided 8-week-long pronunciation training on segmental units for B1-level EFL students to explore whether explicit pronunciation instruction has a superiority over implicit instruction in extensive reading classes. The participants in the control group received implicit pronunciation training in which the language instructor did not directly focus on phonemic properties of English sounds. However, explicit pronunciation training was implemented into the teaching practices in the experimental group. The findings from the study indicate that explicit pronunciation instruction yields better outcomes compared to implicit instruction. According to the author, this could be attributed to the effectiveness of explicit pronunciation training integrated into extensive reading classes. The explicit approach not only enhanced learners' pronunciation skills but also contributed to their overall lexical competence. By focusing on accurate pronunciation of individual segments, learners were able to boost their confidence in using the target language, leading to improved intelligibility as a natural outcome.

Likewise, Lacabex and Gallardo-del-Puerto (2020) compared the effectiveness of explicit phonetic training and implicit exposure to native input for CLIL students' phonological awareness of the occurrence of English schwa in unstressed syllables. Two groups underwent explicit instruction on the incidence of reduced vowels versus full vowels in English disyllabic words, while another group was exposed to native input in their CLIL sessions. A fourth CLIL group with neither explicit intervention nor

native teacher input served as control group. All four groups revealed unawareness of the general pattern of vowel reduction occurrence in unstressed syllables in English prior to intervention. In the posttest, the three experimental groups significantly improved their ability to identify full vowels as incorrect, moreover, the groups receiving explicit instruction exhibited higher gains than the group which was implicitly exposed to native input.

Furthermore, research has also shown the explicit instruction is more advantageous than the implicit approach to boost EFL learners' global intelligibility and speech accuracy. Naeini and Adni (2017) carried out a quasi-experimental study on two homogeneous groups of Iranian high school female students. The experimental group underwent the explicit treatment using a variety of auditory, visual and physical techniques while the control group received oral activities as conventionally instructed in common Iranian high schools. After 15 sessions of the treatment, the results of the posttest rated by an English native speaker indicated that the explicit group outperformed the implicit group concerning speech accuracy and fluency.

Nevertheless, the justification for the efficacy of explicit instruction is not always clear. Lord (2005), for example, drew a conclusion that explicit phonetics instruction is beneficial for L2 Spanish learners. The undergraduate university students in her study improved their pronunciation of different Spanish sounds after an upper-division phonetics course. However, Kissling (2013) raised concerns about Lord's findings due to the absence of a control group in her study. Kissling conducted a similar study but included a control group for implicit instruction. She found that explicit phonetics instruction was not more beneficial than implicit instruction for learners of Spanish. Kissling's study involved participants from different proficiency levels and she proposed that explicit pronunciation instruction might be appropriate for advanced learners, but implicit instruction is more effective for beginners.

Furthermore, Couper (2003) found that intensive explicit instruction on L2 phonological features could help EFL learners reduce some of their pronunciation errors, even when they appear to be fossilized. However, the participants in his study

only retained their pronunciation gains for a short term and their improvements in perception were lost by the end of the semester.

This discrepancy in findings highlights the ongoing debate regarding the effectiveness of explicit versus implicit instruction. It is not rare to find studies that regard implicit instruction as a more advantageous approach. The following session will focus on studies that examine implicit instruction and outline the methodologies used to demonstrate its benefits for learners' pronunciation. It is important to note that shadowing was the primary technique utilized as the implicit treatment in the present research. Therefore, literature discussing the efficacy of imitation or mimicry will also be presented as supporting evidence for implicit instruction.

2.8.2. Implicit instruction

An increasingly researched method in implicit instruction is *shadowing*, an activity where learners imitate a presented speech stimulus as closely and quickly as possible. While shadowing (also referred to as *mirroring* or *tracking*) may not have originated in the language classroom, over the past decade a substantial body of research and books have promoted shadowing as a valuable approach to pronunciation instruction and offered suggestions on how to incorporate it in the classroom (e.g., Foote & Mcdonough, 2017; Hamada, 2018; Hsieh et al., 2013; Kadota, 2019; Yang, 2016). It has been widely acknowledged that learners' exposure to and mimicry of native suprasegmental features is especially beneficial in language learning. Authentic learning materials allow learners to access natural speech with rich prosodic content and full emotional and cognitive expressiveness. Thus, by attempting to replicate the sound patterns of native speakers through specialized techniques, learners are believed to automatically develop connections between ideas in their own minds. This process could not only improve the clarity of their speech but also boost their motivation and confidence in discovering their ability to achieve native-like speech (Yang, 2016, p. 73).

Foote and Mcdonough (2017) investigated the effect of the shadowing method on English learners' ability to imitate speech models as well as their comprehensibility, accentedness, and fluency in extemporaneous talk. 16 participants used iPods to practice shadowing short dialogues for eight weeks. A shadowing task and an extemporaneous speaking task were administered as pre-, mid-, and posttests respectively. The shadowing task was rated for learners' ability to imitate a speech model and the extemporaneous speaking task was rated for comprehensibility, accentedness, and fluency. The recordings were rated by 22 English speakers. Results indicated that the participants improved significantly on all speaking measures other than accentedness. Besides, the interview data collected during the study reflected that participants were largely positive about the mimicry activities, regarding them as an effective way to improve their pronunciation.

These positive results align with research conducted by Vigliano et al. (2016), which likewise revealed that the self-imitation prosodic training could effectively assist L2 learners in improving their rhythmic and prosodic skills, ultimately resulting in improved understanding by native speakers. Similar remarkable results are also detected in Yang's experiment (2016), in which students exhibited a substantial increase in language confidence within a short time, accompanied by significant improvements in pronunciation, presentation skills, and expressiveness of their own ideas and topics.

Despite the growing interest in shadowing, its popularity as a teaching method for pronunciation remains limited in actual practice. This can be attributed, in part, to a lack of sufficient knowledge about the theoretical foundations and justifications underlying this method. For example, researchers who are against this method criticize it as being reminiscent of the audiolingual approach and argue that it is "just vocalized repetitions and only results in meaningless parrot-like practice" (Bovee & Stewart, 2009, p. 20). Additionally, pronunciation has not received ideal emphasis in second or foreign language classrooms, and even less attention has been given to prosody training through imitation.

With regards to other studies that appear to reject the supremacy of explicit approach and favor implicit instruction, Bailey and Brandl (2013) suggested that explicit pronunciation instruction might not be very effective for L2 learners' perceptual development compared to more implicit instruction providing meaning-focused input. Shamiri and Farvardin (2016) conducted interviews with 70 EFL students and found that most of them preferred their errors to be corrected implicitly.

Papachristou (2011) investigated the effects of explicit and implicit instruction on the acquisition of vowels by 15 EFL Greek learners. The results showed that Greek students of both implicit and explicit groups experienced difficulties in producing the target vowels. Regarding vowel duration, it is found that the students in the explicit group achieved some less native-like pronunciation of the vowels compared to the students in the implicit group. She attributed the results to various factors such as the characteristics of the pronunciation features taught and length of instruction.

Similarly, Ruellot (2011) investigated the effects of explicit visual feedback assisted with spectrograms on learners' pronunciation of certain French vowels. The study involved 14 adult learners of French who enrolled in a French phonetics course at an American university. Those participants were randomly assigned to one of two groups—audio feedback condition versus audio-visual feedback condition. Pre- and posttests were administered before and after the instructional treatment. Three native French speakers afterward rated the speech samples produced by the participants from the two tests. The results revealed that participants who were able to see the spectrograms did not improve more than those who were exposed to listening-only practice.

Peltekov (2020) distributed 15 beginning learners of German into three groups (explicit, implicit and control) equally to examine the effects of implicit and explicit instruction on pronunciation respectively. A pretest and posttest design was used to measure learners' improvement in accent and comprehensibility. The results rejected significant differences in participants' progress across the three groups. Specifically, the explicit group tended to perform worse than the implicit group and in some cases

was even outperformed by the control group. However, Peltekov argued that this should not be misinterpreted as evidence of the ineffectiveness of explicit phonetic rules, as "it is possible that students in the explicit group were overwhelmed by the number of rules they had to learn in such a short time (e.g., memorizing a long list of stressed affixes in 10 minutes). They also had less time for practice in comparison to the implicit group" (p. 15). He concluded that some learner variables (e.g., age) might be better predictors of improvement than the type of instruction and participants' individual differences (e.g., motivation level) could significantly affect the extent to which they benefitted from the instruction.

Similarly, Kissling (2013) also made reference to the weight of additional factors to justify the contradictory outcomes associated with explicit versus implicit methods. She pointed out that "It might be the input, practice, and feedback included in pronunciation instruction, rather than the explicit phonetics lessons, that are most facilitative of improvement in pronunciation" (p. 720).

Independently of the methodology favored, an important observation is that most studies have focused on pronunciation instruction of segmental features, while suprasegmental features, especially in terms of word stress and sentence stress instruction, have rarely been investigated. As outlined in the research questions, the present study intends to investigate whether Chinese EFL learners' pronunciation could be improved by a period of training as far as word stress and sentence stress are concerned and whether enhancements vary from different training approaches, i.e., explicit rule explanation vs. implicit imitation.

<div align="center">

CHAPTER III

METHOD

</div>

This chapter firstly outlines the design of the study, which was developed to address the research questions. Subsequently, key information about the participants will be provided and then, the assessment materials and training courses designed for the experiment's objectives will be selectively presented, followed by the questionnaires distributed upon completion of the experiment. Finally, the data analysis tools are explained in order, including acoustic analysis, evaluation for spontaneous speech, and qualitative data analysis. Accordingly, the statistical analysis methods are presented.

3.1. The design

To investigate the efficacy of explicit versus implicit instruction in improving pronunciation, an experiment was conducted involving two groups of participants: the explicit group and the implicit group. The participants were evenly and randomly distributed between these two experimental groups, and both groups attended a pronunciation training course with a focus on the placement of word stress and sentence stress.

Following the same syllabus, however, the teaching approaches and materials varied between the two experimental groups. The explicit group was exposed to the pronunciation rules through different types of sentence stress, namely, sentence-end tonic stress, stress shift for new information, contrastive meaning or emphatic purpose. Simultaneously, the implicit group was provided with the equivalent practice materials accompanied with native speaker recordings to imitate, which the explicit group did not have access to. Thus, the implicit group was expected to infer the underlying rules themselves without explicit instruction from the researcher.

Each group attended eight modules of training, with each lasting approximately 30 minutes. The training sessions occurred twice per week and were conducted remotely via ZOOM due to the COVID restrictions as well as the geographical distance

between the researcher, who served as the course instructor in Spain, and the majority of participants located in China.

The pre- and posttests, both comprising a controlled reading task and a spontaneous speaking task, were administrated for both groups before and after the treatment. The controlled reading data were transcribed and analyzed using the PRAAT software (Boersma & Weenink, 2022), whereas the spontaneous speech data were evaluated by English native speakers in terms of comprehensibility and aceentedness. Additionally, upon completion of the training, a qualitative questionnaire was carried out to collect participants' opinions on their English pronunciation as well as their feedback toward the training. Therefore, both quantitative data and qualitative data were attended to in order to provide a comprehensive analysis of the methodology.

3.2. Participants

Twenty-eight tertiary-level Chinese EFL learners (20F, 8M) aged from 21 to 29 voluntarily participated in the study. All of them had reached the equivalent level of English B1-B2 (CEFR standards) prior to the experiment, as determined by their performance in recognized English proficiency exams (e.g. IELTS, TOEFL, CET6, TEM 4). The majority of participants were tertiary students enrolled in language majors (English & Spanish) or recent graduates residing in China, with the exception of one participant who was born and was studying in Spain majoring in Engineering during the experiment period and another who had previously studied in Spain for a year.

The recruitment of participants was conducted online, with information about the research shared on the Chinese social media platform WeChat as well as several Chinese universities that the researcher had contacts with. All participants willingly enrolled in the pronunciation training course without any form of compensation. During the registration process, they all revealed a desire to enhance pronunciation skills through a casual conversation with the researcher. They recognized their problem of lacking proper rhythm in their speech and expressed the need to learn more about

English prosody. Hence, the aforementioned reasons motivated them to participate in the course.

Afterwards, participants were randomly distributed to the two experimental groups where distinct approaches—explicit or implicit—were implemented. Each group had 14 participants.

3.3. Materials

3.3.1. Pre & posttest

All participants took a pretest prior to the treatment and a similar posttest upon completion of the training. The tests (see Appendix A) consisted of two sections: the first section remained consistent in both tests, involving a controlled reading task of a series of words followed by sentences and dialogues. Meanwhile, the second section, which involved spontaneous speaking tasks using provided picture prompts, exhibited slight variations in topic.

In terms of the selection of words for controlled reading, each word contained a minimum of two syllables so that the stress difference between syllables (especially the difference from the primary stress) could be identified. The chosen target words for this task were: "curi**o**sity", "main**tain**", "a**na**lysis", "satis**fac**tory", "agri**cul**ture", "**a**nalyze", "co**o**perate" and "**i**mmigrant". The syllables in bold indicate the primary stress within each word, although they were not highlighted in the pre- and posttests for participants.

Concerning controlled discourse reading, the materials are separated into four parts covering distinct types of sentence stress—final tonic stress, new information stress, contrastive stress and emphatic stress. Examples are provided in Table 3. The words in bold indicate the tonic stress placement within each sentence, likewise, they were not revealed for participants.

Table 3

Controlled sentence reading tasks in the pre- and posttest

Type of sentence stress	Examples (dialogues and sentences)
Part 1. Final tonic stress	A: Tom lost his **key**. B: I saw it on the **table**. A: Have you **taken** it? B: Yes, I put it in my **bag**. A: Can you **give** it to him? B: Sure. **Where** is he? A: He is playing on the **playground**. B: Ok, I'll meet him after the **class**.
Part 2. Stress shift for new information	A: I can't find my **dress**. B: What **kind** of dress? A: It is a **beach** dress. B: What **color** beach dress? A: It is **white**. White with **stripes**. B: There was a white dress with stripes in the **car**. A: **Which** car? B: The one I **sold**.
Part 3. Contrastive stress	1. He is a **doctor** and I am a **nurse**. 2. I love **playing** badminton rather than **watching** it. 3. Is he **traveling** here? No, he is **living** here. 4. I prefer a **car** to a **bike**. 5. **He** cleaned the room, not **me**. 6. It is **cold** in the morning and **hot** in the afternoon.
Part 4. Emphatic stress	1. A: How was the movie? B: It was **very** boring. 2. A: Why is he not here? B: He is **always** late. 3. A: How is your holiday? B: It is **absolutely** interesting. 4. A: How do you like the view? B: It is **really** pretty. 5. A: Do you like the ice cream? B: Yes, it is **so** tasty. 6. A: Can you go to the meeting? B: I told you I can't go. I feel **terribly** sorry.

As for the spontaneous speech, participants in the pretest were asked to talk about the advantages and disadvantages of living in the city or countryside with questions and pictures provided as prompts. Figure 2 shows the countryside prompt as an example[5]. In the posttest, the same procedure was followed except for the topic being changed to comparing benefits and drawbacks of online and offline shopping. This was done to ensure that speaking prompts were not repetitive, as repetition could potentially influence participants' production. Figure 3 shows the picture prompt used for the posttest.

Figure 2

Example of picture prompt for free-talk in the pretest

---

[5] The city picture is not presented due to copyright protection.

Figure 3

Picture prompt for free-talk in the posttest

3.3.2. Training materials

The pronunciation training modules were originally designed by the researcher, with the goal of instructing Chinese adult EFL learners on the placement of word stress and sentence stress. Even though they followed the same syllabus, the teaching materials varied slightly between the two experimental groups for them to align with the advocated explicit versus implicit teaching methods.

In the explicit group, the pronunciation rules were presented in a clear and direct manner. For instance, students received detailed instructions on when and how to apply different types of sentence stress, including sentence-end tonic stress, stress shift for conveying new information, and stress for contrastive or emphatic purposes. Figure 4 provides an overview of the teaching syllabus, highlighting the specific topics covered. The participants were guided through a progression, starting from distinguishing syllables and various levels of word stress to mastering the nuances of different sentence stress patterns.

Figure 4

Teaching syllabus designed for the explicit group

Furthermore, Figure 5 provides an illustrative example of the teaching materials exclusively tailored for the explicit group. The complete compilation of the entire explicit course materials can be found in Appendix B.

Figure 5

Example of teaching materials for the explicit group

Simultaneously, the implicit group was instructed through imitation practice. Participants were provided with the equivalent practice materials utilized in the course, complemented by recordings from two native speakers (one from the US and the other from the UK) for imitation, which the explicit group did not have access to. The native speakers, despite being encouraged to produce the utterances in a natural manner, were also briefed on the specific teaching focus regarding different types of stress before making recordings, so they were particularly attuned to the target positions of prominence within each sentence. This helped regulate the placement of stress and minimize any potential contradiction between native speech and the target rules. The implicit group, in the absense of explicit guidance from the researcher, was expected to infer the underlying rules themselves by imitating the audio samples. Figure 6 provides an overview of the teaching syllabus, where the keyword *stress* was not explicitly revealed. Instead of being directly taught, all relevant information was conveyed in a manner that attracted participants' attention to the target research focus.

Figure 6

Teaching syllabus designed for the implicit group
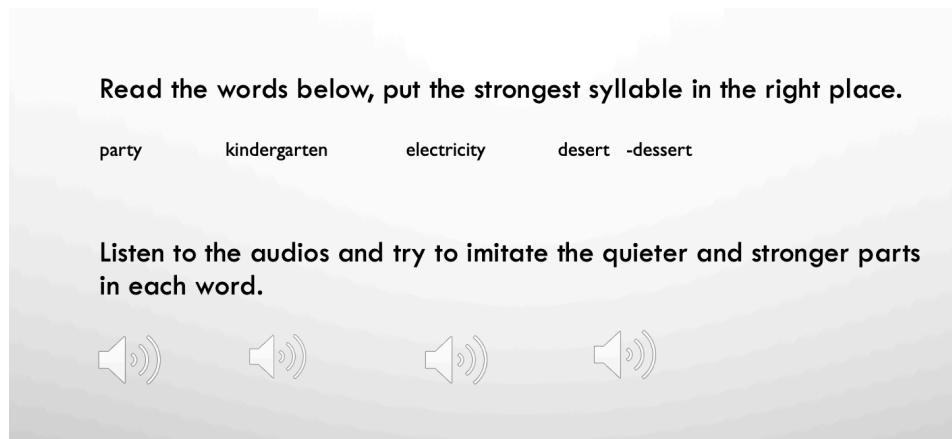
Figures 7 and 8 show screenshots from the course, exemplifying the teaching materials tailored for the implicit group. A comprehensive collection of all course materials can be found in Appendix C.

Figure 7

Example of teaching materials for word stress in the implicit group

Read the words below, put the strongest syllable in the right place.

party   kindergarten   electricity   desert -dessert

Listen to the audios and try to imitate the quieter and stronger parts in each word.
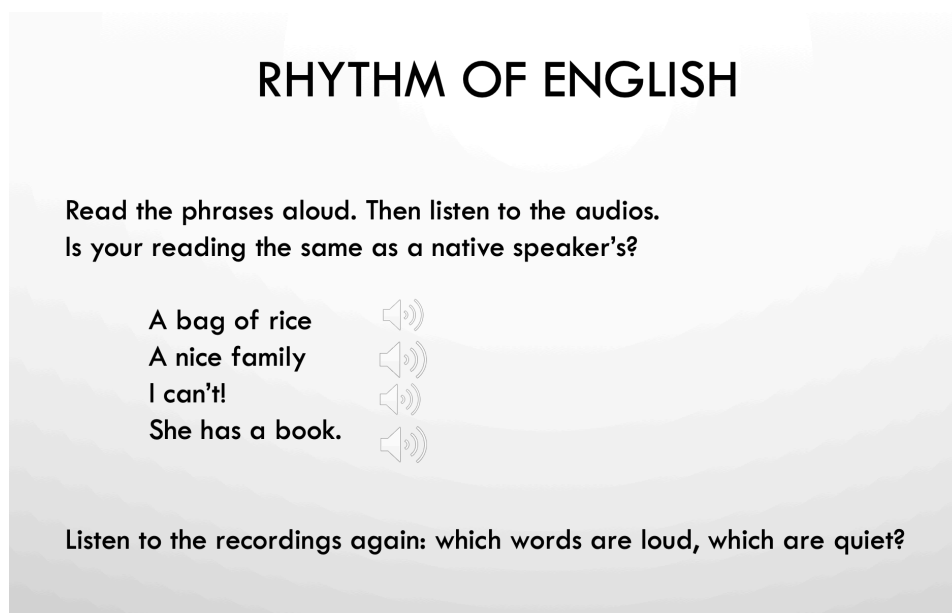
Figure 8

Example of teaching materials for sentence stress in the implicit group

RHYTHM OF ENGLISH

Read the phrases aloud. Then listen to the audios.
Is your reading the same as a native speaker's?

A bag of rice
A nice family
I can't!
She has a book.

Listen to the recordings again: which words are loud, which are quiet?

3.3.3. Qualitative questionnaire

Upon completion of the training course, the participants were asked to respond to a qualitative questionnaire that incorporated both 5-point Likert scale questions and open-ended questions. The questionnaire encompassed two sections. The first part remained the same for the two groups of trainees, aiming to collect their background

information, such as name, age, sex, and level of English, as well as their opinions on the importance of English pronunciation and the effectiveness of the course they took. In the second part, specific questions were tailored exclusively for each experimental group to gather their feedback on the training sessions. For example, the explicit group was asked about the level of difficulty they experienced in understanding the rules explained in class, how useful they found the training to be and whether they would prefer to use native speaker audios as supplementary resources for imitation. On the other hand, the implicit group was questioned about the extent to which they felt the imitation practice improved their pronunciation and whether they encountered difficulties in following the imitation exercises etc. The complete questionnaire can be found in Appendix D.

3.4. Data collection and analysis instruments

In terms of the pre-test and post-test data collection and analysis, only seven participants from the explicit group and ten from the implicit group completed 90% or above of the training course. Their pre and post-test recordings were then used for acoustic and statistical analysis, followed by human judges' evaluation. Given the restrictions imposed by the remote research conditions, the recordings were done on the participants' own mobile devices in a required noise-free environment then transferred to the researchers by email. Due to time limitations, a decision was made to acoustically analyze only recordings of words and contrastive stress sentences from the pre- and posttest, along with judges' evaluation of the spontaneous speech. Some participants' reponses from the satisfaction surveys will also be selectively presented and discussed. The following sub-sections describe the different approaches adopted to analyze the data.

3.4.1. Acoustic analyses of the controlled-reading data

A part of the controlled-reading recordings—word reading and contrastive stress sentencess—was transcribed and annotated on the PRAAT speech analysis software

48

for acoustic analysis. As far as word stress is concerned, each syllable in the word was identified in the acoustic signal and segmented manually based on audio perception and spectrum. A reading of the intensity peak, pitch average, and duration was then obtained through a PRAAT script (see Appendix E). In Figure 9, an example of the word "maintain" is shown. The yellow lines depict the intensity values of the syllables within the word and the blue curves display pitch values. Notably, there is a discernible discrepancy in amplitude peaks between the stressed syllable "tain" and the unstressed syllable "main".

Figure 9

The waveform, spectrum and annotation of word "maintain"
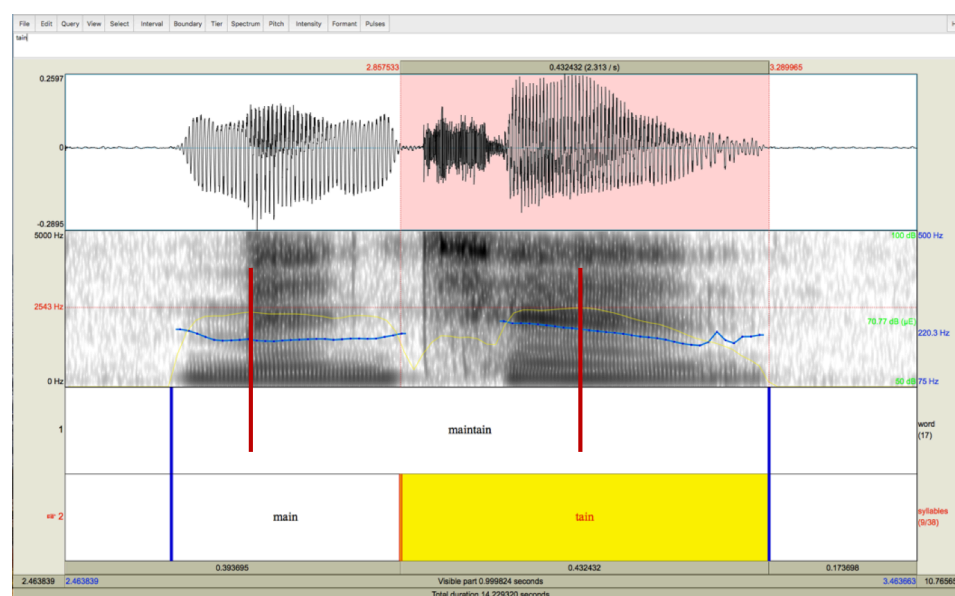
Note. The ruby vertical lines represent the location of the intensity peak in each syllable.

As far as intensity is concerned, the target scores were obtained by subtracting the intensity peak value of an unstressed syllable from that of the primary-stress syllable. Subsequently, these intensity difference scores were compared across each participant between the pretest and the posttest, as well as between the explicit and the implicit group. Table 4 illustrates the subtraction process used to calculate the intensity difference in the word "analyze". Here, the stressed syllable is "a" and the other selected

syllable for the subtraction is "lyze"[6]. For the sake of time efficiency, within each word only one subtraction was conducted. The final column shows the difference in amplitude for the selected syllables. Scores from additional participants in both the pre-test and post-test stages will be provided in the Results chapter.

Table 4

Intensity peaks of syllables in "analyze" and intensity difference between syllables

Word	Unstressed syllable intensity peak (dB)	Stressed syllable intensity peak (dB)	Difference in amplitude(dB)
analyze	**Sy3 "lyze"**	**Sy1 "a"**	**Sy1-3**
pre	78.26	76.95	-1.31
post	76.83	78.43	1.60

The same calculation procedure was applied to ascertain the disparity in average F0 values between the pre- and posttest, as shown in Table 5.

Table 5

F0 average of syllables in "analyze" and F0 difference between syllables

Word	Unstressed syllable F0 average (Hz)	Stressed syllable F0 average (Hz)	Difference in F0 (Hz)
analyze	**Sy3 "lyze"**	**Sy1 "a"**	**Sy1-3**
pre	218.93	235.00	16.06
post	202.13	240.10	37.97

---

[6] Certain words contain different levels of stress, but selected syllables are consistently subtracted from the primary stress, which results in discernible differences in intensity, F0, and duration.

As for syllable duration measurements, the duration of each syllable was identified and obtained from the segmentation previously done for intensity and F0. For instance, as presented in Figure 10, the stressed syllable "tain" had a longer duration, compared to the less stressed syllable "main". The respective durations of the two syllables are 0.43s and 0.27s.

Figure 10

An example showing durations of syllables in "maintain"
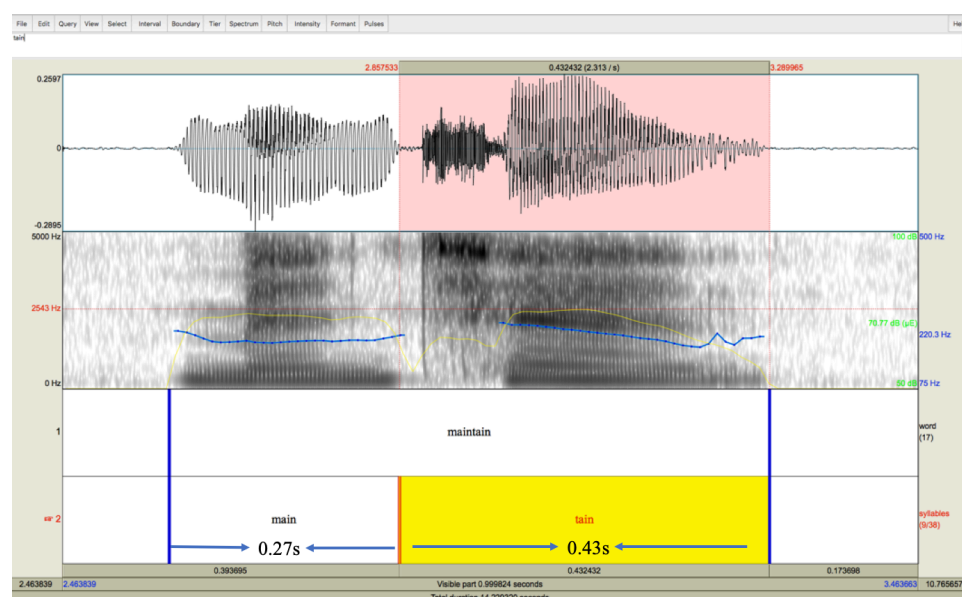
Subsequently, the division was employed to calculate proportion ratios across syllables, particularly, the percentage of the unstressed syllable duration divided by the primary-stress/nuclear syllable within the same word. The smaller the proportion, the higher difference there is between the nuclear syllable and the unstressed one. An example of this calculation is provided in Table 6.

51

Table 6

Duration of syllables in "maintain" and the obtained ratios in the pre and post

Word	Unstressed syllable's duration (s)	Stressed syllable's duration (s)	Percentage in duration
maintain	**Sy1 "main"**	**Sy2 "tain"**	**Sy1/Sy2**
pre	0.31	0.42	73.80%
post	0.27	0.43	62.79%

In terms of contrastive stress analysis at the sentence level, a slightly different procedure was followed. Instead of segmenting individual syllables, each word in the sample sentences was segmented, and readings were taken for intensity peak, F0 average, and duration. Similar to the word stress analysis, relative differences in acoustic cues were calculated between the target tonic stress and another selected stressed word within the same sentence. Figure 11 illustrates the segmentation of the sentence "He is a doctor and I am a nurse" on PRAAT for measuring acoustic cues, while Table 7 provides an example of intensity reading and calculation in the pretest and posttest from one of the participants. It was assumed that the most semantically important words in the sentence were "doctor" and "nurse", which also carry contrastive information, so they were treated as tonic stress. Thus, the acoustic values of other words of the same sentence were calculated with these two. In terms of F0 and duration values obtained from the contrastive sentences, the same procedure was followed as in word stress analysis.

Figure 11

An example of sentence segmentation and acoustic patterns
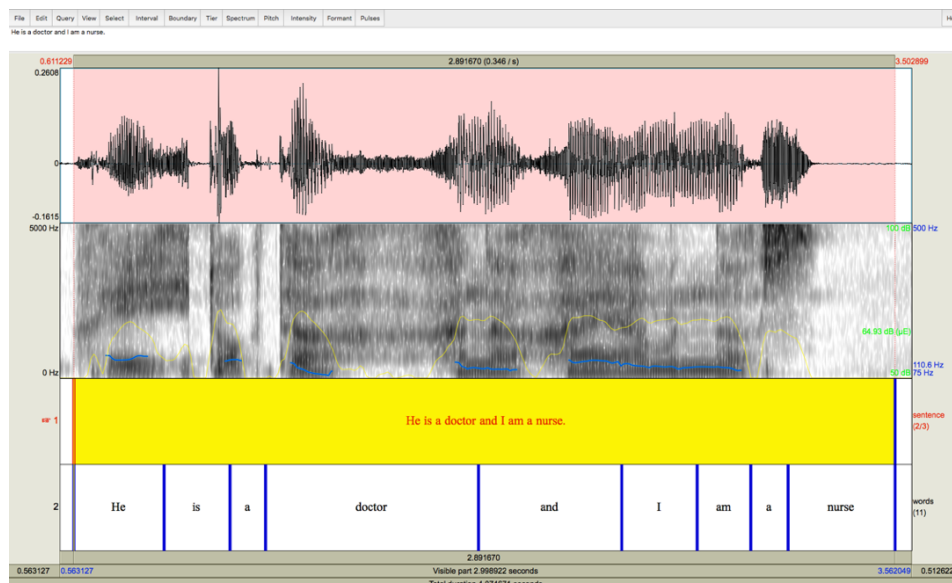
Table 7

Intensity peaks of selected words in a sentence and calculated value differences

Intensity peak values of selected words (dB) in "He is a doctor and I am a nurse"				
Selected words	**He**	**doctor**	**I**	**nurse**
	68.17	71.89	70.16	69.71
Subtraction values (dB)	**Doctor** - he	3.72	**nurse** - I	-0.45

Finally, the difference scores in intensity and F0, along with the percentage rates of duration, in the pre- and posttests, were submitted to statistical analyses. In order to test the significance of the differences between the pre- and posttest as well as differences across groups, a series of two-way repeated measures ANOVA tests were performed. The statistical calculation included two independent variables, i.e., time and group, and the independent variable, namely, the calculated difference of acoustic values obtained

in the pre- and posttests. Following the two-way repeated measures ANOVA tests, Bonferroni post-hoc tests were subsequently conducted for multiple comparisons, especially when a statistical significance was deteact in the ANOVA test. All the tests were performed on the statistical analysis application JASP. The level of statistical significance for all tests was set at p < .05.

It was hypothesized that if participants were aware that stressed syllables exhibit greater intensity, higher pitch, and longer duration, both the amplitude and F0 differences between stressed and unstressed syllables would increase from the pretest to the posttest. However, the percentage of unstressed syllable duration divided by the stressed would diminish.

3.4.2. Collection and analyses of spontaneous speech data

Regarding the spontaneous production data, each participant from the two groups generated approximately one minute of free speech in the pre- and posttest respectively. In order to examine the controlled reading and spontaneous data of the same participants, the analyzed spontaneous productions were obtained from the same 7 participants in the explicit group and 10 in the implicit group, whose controlled reading data were previously analyzed acoustically. In total, 17 participants produced 34 free-talk audio items, all of which were randomized and uploaded into a Google Form along with a 7-point Likert scale evaluation form for comprehensibility and accentedness. Subsequently, 12 native English speakers were recruited as evaluators to assess these audio files, with their primary focus on the following two aspects—speech comprehensibility (the ease of understanding the speaker) and accentedness (the degree of the speaker's foreign accent), as depicted in Figure 12. In the evaluation of comprehensibility, in order to avoid ambiguity, raters were provided with detailed 7-point scale anchor indications: anchor 1 *not comprehensible at all* means "you find it impossible to tell what is being said", and anchor 7 *highly comprehensible* refers to the fact that "you have to make no effort at all to understand what is being said".

Figure 12

Example capture of the assessment form for native speaker judges

To ensure that the judges' perception of the speakers' prosody production remained uninfluenced by disfluency, proactive measures were taken to edit the audio samples in advance. This editing involved removing noticeable pauses and mitigating any unpleasant echoes or background noise. In addition, on the evaluation form, raters were explicitly reminded to solely concentrate on comprehensibility and foreign accent of the speech, without considering use of vocabulary, grammar, or speech fluency.

All of the invited raters were English native speakers, adults, living in Spain at the time of evaluation except one in Scotland. Table 8 shows the nationality and the occupations of these raters.

Table 8

Raters' demographic background information

Raters	Nationality	Occupations
1	Canadian	University teacher & researcher
2	English	University teacher & researcher
3	English	University teacher & researcher
4	Australian	University teacher & translater
5	American	ESL teacher
6	English	University teacher & researcher
7	English	University teacher & translater
8	Scotish	University teacher & researcher
9	English	University teacher & researcher
10	American	ESL teacher
11	English	University teacher & researcher
12	American	ESL teacher

Regarding the statistical analysis methods for the spontaneous data, first of all, a reliability analysis using an Intra-Class Correlation Coefficient (ICC) was performed to determine the level of agreement between the 12 judges for the evaluation task. Subsequently, a series of two-way repeated measures ANOVA tests were conducted to investigate if the training led to any improvement in free speech for each group. Simultaneously, the average scores in comprehensibility and accentedness were calculated for each test and group. Furthermore, where significance was detected in ANOVA, post-hoc tests were performed for multiple comparisons.

3.4.3. Analysis of qualitative survey data

The qualitative data gathered from the post-training questionnaires will be presented selectively. Where applicable, t tests were also performed to explore

statistical differences in responses across groups. Results of all aforementioned tests

can be found in the subsequent chapter.

CHAPTER IV
RESULTS

The present chapter is divided into three main sections. Section 4.1 presents the results of the acoustic analysis of the data covering changes in intensity, F0 and duration from the pre- to the posttest; it also reports the results of a series of two-way repeated measures ANOVA tests conducted to examine differences between the groups at both test stages. Post-hoc tests were subsequently performed for multiple comparisons. For all the tests, the level of statistical significance was set at $p < .05$.

Section 4.2 shows the results from the 12 native-speaking judges who evaluated the participants' spontaneous speech for general comprehensibility and accentedness. This section includes a reliability analysis of ICC, rating means for each test and group, as well as the two-way repeated measures ANOVA test results exploring time and group effects.

Section 4.3 selectively presents the results obtained from the qualitative questionnaire. For the same Likert scale questions distributed to the two groups for collecting participants' views towards pronunciation learning and feedback to the training, a series of t tests were conducted to compare the difference between the two groups' responses.

4.1. Results of the acoustic analyses of word stress

4.1.1. Analysis of intensity peak for word stress

Table 9 shows the mean intensity peak values obtained from the subtraction process of the syllables as far as word stress is concerned. The first column refers to the position of the unstressed syllables which were subtracted from the stressed syllables. The numbers in the two columns under each group present the mean difference obtained from the subtraction generated by all participants at the two test stages. If we compare the mean scores of the values of both groups obtained in the pre- and posttests, we observe that all of the figures in the posttest were larger than those in the pretest except

59

for the three sets in red under the implicit group. Regarding the negative figures, they indicate that some participants produced unstressed syllables with higher intensity than the primary stressed syllable.

Table 9

Mean differences of word intensity peaks in two groups at the pre- and posttest

	Explicit group		Implicit group	
Syllables for subtraction	Mean difference in Pre	Mean difference in Post	Mean difference in Pre	Mean difference Post
Curiosity				
Sy3-Sy1	-2.83	-1.15	-1.73	-1.21
Maintain				
Sy2-Sy1	-1.38	0.81	0.16	0.42
Analysis				
Sy2-Sy1	0.05	3.33	-0.54	1.48
Satisfactory				
Sy3-Sy1	-0.42	1.70	0.15	-0.83
Agriculture				
Sy3-Sy1	-1.75	-0.66	-1.15	0.17
Analyze				
Sy1-Sy3	1.33	2.77	2.88	1.54
Cooperate				
Sy2-Sy1	2.27	5.05	4.42	3.45
Immigrant				
Sy1-Sy3	0.50	0.90	0.21	1.60

In order to test the significance of the differences between the pre- and posttest as well as differences across groups, two-way repeated measures ANOVAs were employed. Each participant produced reading data for eight words, and only one subtraction was done for each word, as shown in Table 9. The recordings from the 7 participants in the explicit group and 10 from the implicit group were considered valid for analysis. That resulted in 56 (7 participants *8 words) valid tokens for the explicit group and 80 (10 participants *8 words) tokens for the implicit group. Table 10 presents the word intensity difference for each group between the two time points in terms of

means and standard deviations. Both groups show an increase in their values in the posttest compared with the pretest. The improvement was more pronounced in the explicit group: $M = -0.28$, $SD = 2.74$ in the pretest and $M = 1.62$, $SD = 3.89$ in the posttest. The implicit group, on the other hand, obtained $M = 0.55$, $SD = 3.72$ in the pretest, and $M = 0.83$, $SD = 3.20$ in the post-test conditions. Figure 13 illustrates these differences in a plot.

Table 10

Descriptive statistics for word intensity peaks in two groups at the pre- and posttest

		M	SD	N
Explicit	Pretest	−0.28	2.74	56
	Posttest	1.62	3.89	56
Implicit	Pretest	0.55	3.72	80
	Posttest	0.83	3.20	80

Figure 13

Plot of mean differences of word intensity in two groups at the pre- and posttest
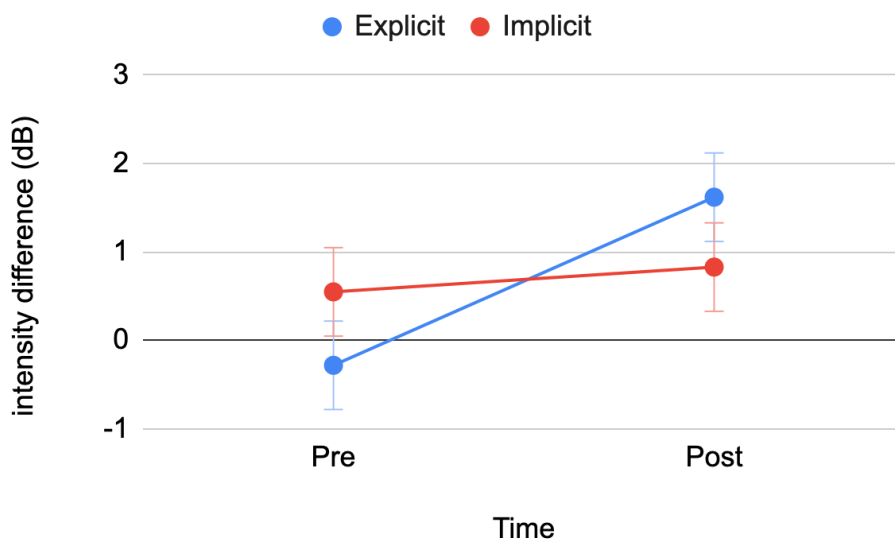
The two-way repeated measures ANOVA test reveals that both the time variable and the time*group interaction were significant within-subject effects, as shown by the p values in Table 11 and illustrated in Figure 13. However, in terms of between-subject

effects, i.e., for the group variable, significance was not detected across groups, $p = 0.978$, as shown in Table 12.

Table 11

ANOVA test results for word intensity - within subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Time	78.108	1	78.108	10.950	0.001
Time *Group	43.413	1	43.416	6.087	0.015
Residuals	955.829	134	7.133		

Note. Type III Sum of Squares

Table 12

ANOVA test results for word intensity - between subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Group	0.012	1	0.012	7.388×10^{-4}	0.978
Residuals	2203.362	134	16.443		

Note. Type III Sum of Squares

To further investigate the specific points of significance, a Bonferroni post-hoc test for pairwise comparison was carried out, the results of which are comprehensively presented in Table 13. The most important pairwise comparisons between the two tests and across the two groups are highlighted in red for emphasis. Initially, there was no notable difference in the pretest between the two groups ($t = -1.38$, $p = 0.507$), indicating that both groups entered the experiment at a similar level. However, the explicit group witnessed a significant difference between the pre- and posttests ($t = -3.766$, $p < .05$), whereas the implicit group did not ($t = -0.656$, $p = 0.513$). At the posttest, once again, there was no significant disparity between the two groups ($t = 1.334$, $p = 0.507$). These results are behind the previously shown significant time*group

interaction, where one group demonstrated significant improvement while the other did not. Finally, neither group exhibited significant differences in either pre- or posttest, which aligns with the results presented in Table 12, where the time variable was a significant factor while the group variable was a non-significant factor.

Table 13

Post-hoc results for word intensity in two groups at the pre- and posttest

		Mean Difference	*SE*	*t*	Pholm
Explicit, Pre	Implicit, Pre	−0.825	0.598	−1.380	0.507
	Explicit, Post	−1.901	0.505	−3.766	0.001
	Implicit, Post	−1.102	0.598	−1.843	0.333
Implicit, Pre	Explicit, Post	−1.075	0.598	−1.797	0.333
	Implicit, Post	−0.277	0.422	−0.656	0.513
Explicit, Post	Implicit, Post	0.798	0.598	1.334	0.507

Note. P-value adjusted for comparing a family of 6

4.1.2. Analysis of F0 average for word stress

In terms of F0 analysis, Table 14 shows the means of F0 subtraction values in each word, gathered from all participants in the pretest and the posttest from each group. As in the intensity analysis, most of the values in the posttest are larger than the pretest except for a few marked in red.

Table 14

Mean differences of word F0 average in two groups at the pre- and posttest

Syllables for subtraction	Explicit group		Implicit group	
	Mean difference in Pre	Mean difference in Post	Mean difference in Pre	Mean difference Post
Curiosity				
Sy3-Sy1	-17.46	-1.40	-8.50	-14.25
Maintain				
Sy2-Sy1	-5.66	6.80	-10.19	-0.62
Analysis				
Sy2-Sy1	8.24	17.52	-1.52	5.92
Satisfactory				
Sy3-Sy1	9.84	17.87	5.38	16.27
Agriculture				
Sy3-Sy1	-8.04	-10.13	0.49	15.64
Analyze				
Sy1-Sy3	47.13	66.59	65.23	73.48
Cooperate				
Sy2-Sy1	5.33	5.66	4.51	27.50
Immigrant				
Sy1-Sy3	31.19	71.25	60.95	79.00

Table 15 displays the descriptive statistics from the two-way repeated measures ANOVA test taking all words into account. In the same vein, both groups have shown a higher F0 difference in the posttest, however, the explicit group reveals a bigger difference between pre- and posttest (Explicit: pre $M = 8.82$, $SD = 29.97$ vs. post $M = 21.60$, $SD = 36.69$; Implicit: pre $M = 14.54$, $SD = 37.26$ vs. post $M = 23.68$, $SD = 39.57$). Figure 14 illustrates the changes from the pretest to the posttest in each group in a plot.

Table 15

Descriptive statistics for word F0 average in two groups at the pre- and posttest

		M	SD	N
Explicit	Pretest	8.82	29.97	56
	Posttest	21.60	36.69	56
Implicit	Pretest	14.54	37.26	80
	Posttest	23.68	39.57	80

Figure 14

Plot of mean differences of word F0 average in two groups at the pre- and posttest
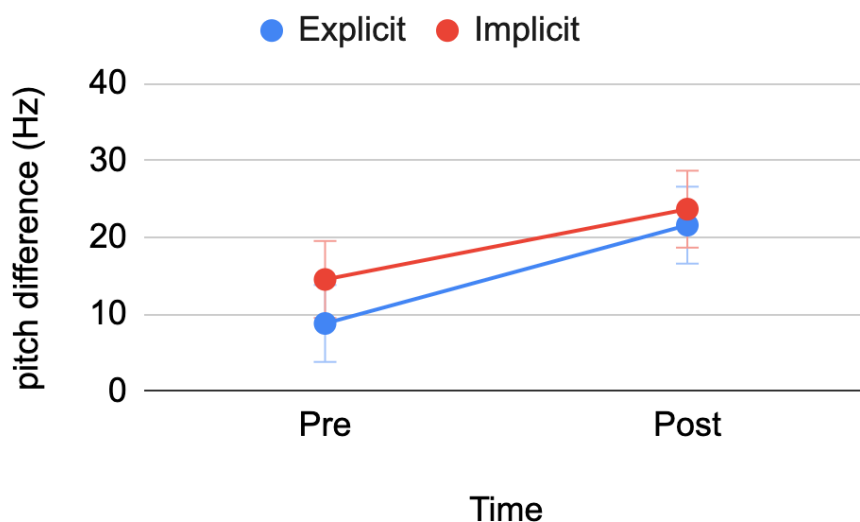
Table 16 and 17 present the statistics of the two-way repeated measures ANOVA test, which reveals that, for the within-subject effects, only the time variable played a significant role ($p < .001$), whereas the time*group interaction was non-significant ($p = 0.469$). Additionally, no significance was found for the between-subject effects ($p = 0.506$).

Table 16

ANOVA test results for word F0 average - within subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Time	7909.064	1	7909.064	19.115	< .001
Time *Group	218.350	1	218.350	0.528	0.469
Residuals	55443.521	134	413.758		

Note. Type III Sum of Squares

Table 17

ANOVA test results for word F0 average - between subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Group	1002.337	1	1002.337	0.446	0.506
Residuals	301342.350	134	2248.824		

Note. Type III Sum of Squares

Slightly different from the statistics for intensity, the post-hoc tests in Table 18 indicate a significant improvement in F0 difference between pre- and posttests for both groups (Explicit: $t = -3.324$, $p < .05$; Implicit: $t = -2.841$, $p < .05$). Still, neither group exhibited a significant difference at either pre or posttest (Explicit pre versus Implicit pre: $t = -0.90$, $p = 0.805$; Explicit post versus Implicit post: $t = -0.327$, $p = 0.805$). These findings support the results drawn from the ANOVA tests, suggesting that the time variable played a significant role in both groups, while no significant group effect was identified. As both groups demonstrated substantial improvement simultaneously, no interaction between the groups was detected, as shown in Figure 14.

Table 18

Post-hoc results for word F0 average in two groups at the pre- and posttest

		Mean Difference	SE	t	Pholm
Explicit, Pre	Implicit, Pre	−5.721	6.357	−0.900	0.805
	Explicit, Post	−12.777	3.844	−3.324	0.007
	Implicit, Post	−14.857	6.357	−2.337	0.082
Implicit, Pre	Explicit, Post	−7.056	6.357	−1.110	0.805
	Implicit, Post	−9.136	3.216	−2.841	0.026
Explicit, Post	Implicit, Post	−2.080	6.357	−0.327	0.805

Note. P-value adjusted for comparing a family of 6

4.1.3. Analysis of duration for word stress

As far as duration measurements are concerned, the percentage was obtained from the duration of the unstressed syllable divided by that of the primary stressed syllable within the same word. The same syllables chosen for intensity and F0 analysis were used to calculate the division. Table 19 shows the means of duration percentages produced by all subjects for each word in the pretest and posttest across the two groups. As expected, the percentage values in most cases decreased in the posttest compared with the pretest, except those marked in red. Such a drop means that either the stressed syllable was pronunced with longer duration or the unstressed syllable was shorter. However, it is also important to note that many percentage figures exceed 100%, which indicates that the unstressed syllables in those words were actually pronounced longer than the stressed syllables.

Table 19

Means of duration ratios within each word at the pre- and posttest in two groups

	Explicit group		Implicit group	
Syllables for division	Percentage in Pre	Percentage in Post	Percentage in Pre	Percentage in Post
Curiosity				
Sy1/Sy3	136.83%	125.79%	155.31%	141.49%
Maintain				
Sy1/Sy2	63.12%	61.69%	71.79%	64.87%
Analysis				
Sy1/Sy2	110.40%	95.68%	96.22%	79.12%
Satisfactory				
Sy1/Sy3	104.66%	99.73%	101.74%	94.60%
Agriculture				
Sy1/Sy3	68.07%	69.48%	73.13%	69.77%
Analyze				
Sy3/Sy1	261.04%	287.94%	276.72%	329.12%
Cooperate				
Sy1/Sy2	148.47%	126.60%	127.07%	103.70%
Immigrant				
Sy3/Sy1	417.26%	344.23%	310.90	386.84%

Table 20 presents the descriptive statistics for the duration ratios at the two test points in each group, focusing on means and standard deviations. It is evident that both groups experienced a reduction in ratios between the pretest and the posttest, i.e., in the explicit group: $M =163.7\%$, $SD = 131.2$ in the pretest and $M =151.4\%$, $SD =107.3$ in the posttest, whereas the implicit group, on the other hand, obtained $M =161.1\%$, $SD = 115.0$ in the pretest, and $M = 149.2\%$, $SD =112.2$ in the posttest. The percentage decrease from the pretest to the posttest was quite similar in both groups, with the explicit group experiencing a 13% reduction and the implicit group showing a 12% decrease.

Table 20

Descriptive statistics for word duration ratios in two groups at the pre- and posttest

		M	SD	N
Explicit	Pretest	163.7%	131.2	56
	Posttest	151.4%	107.3	56
Implicit	Pretest	161.1%	115.0	80
	Posttest	149.2%	112.2	80

Figure 15 visually illustrates the changes in duration percentage from the pretest to the posttest across both groups in a plot. The change observed in each group followed a strikingly similar pattern. Both groups showed a drop in the percentage of duration; however, it is noteworthy that the disparity between the two groups is relatively small.

Figure 15

Plot of changes in word duration ratios in two groups at the pre- and posttest
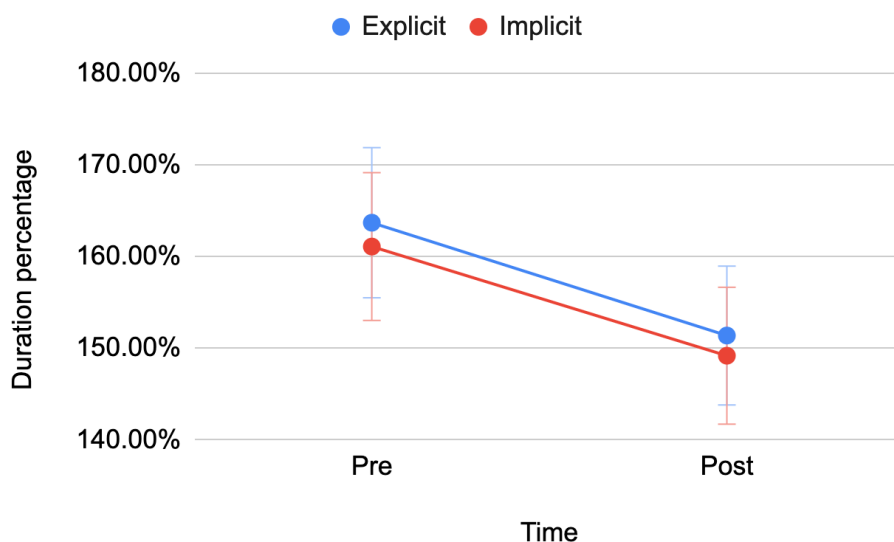
In order to ascertain the significance of these observations, a two-way repeated measures ANOVA test was conducted. The time variable once again yielded

significance, $p < .001$, as shown in Table 21, whereas there was no significant between-subject effect across the groups, $p = 0.901$, as evidenced in Table 22.

Table 21

ANOVA test results for word duration - within subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Time	0.968	1	0.968	4.334	< .001
Time *Group	3.137×10^{-4}	1	3.137×10^{-4}	0.001	0.970
Residuals	29.918	134	0.223		

Note. Type III Sum of Squares

Table 22

ANOVA test results for word duration - between subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Group	0.038	1	0.038	0.015	0.901
Residuals	332.005	134	2.478		

Note. Type III Sum of Squares

To further find out where significance was generated within the groups, post-hoc tests were carried out and the statistics are presented in Table 23. As predicted, there was no significant difference between the two groups in either the pre- or posttest. Unexpectedly, within each of the two groups, no significant reduction in duration was observed from the pre- to the posttest. This might appear contradictory to the significant effect of time variable in the ANOVA results, but it can be interpreted as indicating that pairwise comparisons in the post hoc-test do not reflect this significance, possibly due to the relatively small sample sizes.

Table 23

Post-hoc results for word duration ratios in two groups at the pre- and posttest

		Mean Difference	*SE*	*t*	Pholm
Explicit, Pre	Implicit, Pre	8.1%	17.4%	46.5%	1.000
	Explicit, Post	12.2%	9.1%	134.0%	0.912
	Implicit, Post	20.1%	17.3%	115.9%	0.992
Implicit, Pre	Explicit, Post	4.1%	17.3%	24.0%	1.000
	Implicit, Post	12%	7.6%	158.1%	0.698
Explicit, Post	Implicit, Post	7.8%	17.4%	45.1%	1.000

Note. P-value adjusted for comparing a family of 6

4.2. Results of the acoustic analyses of contrastive sentence stress

4.2.1. Analysis of intensity peak for contrastive sentence stress

Regarding contrastive sentence stress, Table 24 shows mean values of intensity peak obtained from the subtraction process of the selected words within a sentence. The first column refers to the positions of the unstressed words subtracted from the words carrying contrastive stress. Within each sentence, two subtraction processes between stressed and unstressed words were conducted. For example, in the sentence "He is a doctor and I am a nurse", two target stressed words "doctor" and "nurse" were considered, leading to two subtractions that involved these words as well as another two unstressed words "he" and "I", respectively.

The two columns in each group present the mean differences obtained from the subtraction process carried out from all participants at the two test stages. When we compare the mean scores of both groups between the pre and posttests, it is observed that all the figures in the posttest are larger than those in the pretest, despite the presence of some negative figures.

Table 24

Mean differences of intensity peaks for sentence stress in two groups at the pre- and posttest

	Explicit group		Implicit group	
Words for subtraction	Mean difference in Pre	Mean difference in Post	Mean difference in Pre	Mean difference Post
S1				
doctor – he	-2.26	1.96	-1.40	2.54
nurse – I	0.27	0.94	-1.26	0.42
S2				
playing – badminton	1.41	2.34	1.36	2.57
watching – badminton	-2.46	1.52	-0.69	0.13
S3				
traveling – here	0.73	3.34	1.07	2.41
living – here	4.19	7.78	5.10	7.13
S4				
car – prefer	-2.11	-1.51	-2.98	-1.86
bike – prefer	-3.19	-2.02	-5.45	-1.84
S5				
he – cleaned	-0.40	1.63	-2.35	-0.22
me – not	-0.59	1.38	-2.59	-0.10
S6				
cold – morning	-0.04	2.45	0.31	3.41
hot – afternoon	1.26	2.05	2.33	3.06

In order to test the significance of the differences between the pre- and posttest for each participant as well as differences across groups, a two-way repeated measures ANOVA test again was performed. Since each sentence involves two subtractions and every participant was recorded reading 6 sentences, 12 subtraction values were obtained from each participant. Therefore, in total, 7 participants from the explicit group generated 84 valid tokens, and 10 participants from the implicit group produced 120 tokens. Table 25 presents the means of the gained intensity difference for each group between the two time points and the corresponding standard deviations. A greater contrast was seen in the explicit group (pre $M = -0.29$, $SD = 3.06$ vs. post $M = 1.65$, SD

= 4.02) than the implicit group (pre $M = -0.13$, $SD = 4.56$ vs. post $M = 1.01$, $SD = 4.48$).

Figure 16 presents the plot of mean intensity differences of contrastive sentence stress

between the pre- and posttests in the two groups.

Table 25

Descriptive statistics for sentence stress in two groups at the pre- and posttest

		M	*SD*	*N*
Explicit	Pretest	–0.29	3.06	84
	Posttest	1.65	4.02	84
Implicit	Pretest	–0.13	4.56	120
	Posttest	1.01	4.48	120

Figure 16

Plot of mean intensity differences for sentence stress in two groups at the pre- and

posttest
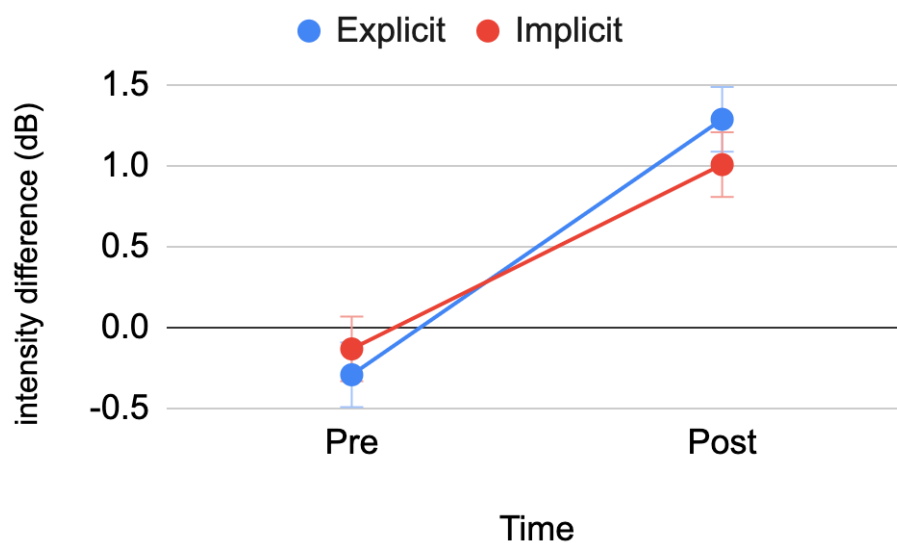
As indicated by the statistics of within-subject effects and between-subject effects

in Table 26 and Table 27 respectively, only the time factor showed significance, *p*

< .001. Neither the group effect nor the time*group interaction displayed any significant

results.

Table 26

ANOVA test results of intensity differences for sentence stress – within subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Time	233.010	1	233.010	29.051	< .001
Time *Group	15.657	1	15.657	1.952	0.164
Residuals	1620.181	202	8.021		

Note. Type III Sum of Squares

Table 27

ANOVA test results of intensity differences for sentence stress – between subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Group	5.585	1	5.585	0.210	0.647
Residuals	5361.345	202	26.541		

Note. Type III Sum of Squares

Seen from the post-hoc test results in Table 28, the improvement was significant in each group between pre and post (Explicit: $t = -4.425$, $p < .05$; Implicit: $t = -3.111$, $p < .05$). There was no noticeable disparity across groups at the either the pretest ($p = 0.787$) or the posttest ($p = 0.566$).

Table 28

Post-hoc results of intensity for sentence stress in two groups at the pre- and posttest

		Mean Difference	SE	t	Pholm
Explicit, Pre	Implicit, Pre	−0.027	−1.730	−0.271	0.787
	Explicit, Post	−1.934	−3.098	−4.425	< .001
	Implicit, Post	−1.298	−2.868	−2.194	0.087
Implicit, Pre	Explicit, Post	−1.773	−3.343	−2.998	0.012
	Implicit, Post	−1.137	−2.112	−3.111	0.011
Explicit, Post	Implicit, Post	0.636	−0.934	1.075	0.566

Note. P-value adjusted for comparing a family of 6

4.2.2 Analysis of F0 average for contrastive sentence stress

Examining the pitch statistics pertaining to contrastive sentence stress, as indicated by the descriptive statistics in Table 29, it is apparent that the explicit group exhibited a bigger difference compared to the implicit group (Explicit pre: $M = -3.27$, $SD = 41.36$ vs. Explicit post $M = 13.61$, $SD = 64.00$; Implicit pre $M = -1.84$, $SD = 44.90$ vs. Implicit post $M = 9.44$, $SD = 47.15$). Figure 17 displays a pattern similar to the intensity results presented earlier.

Table 29

Descriptive statistics of F0 average for sentence stress in two groups at the pre- and posttest

		Mean	SD	N
Explicit	Pretest	−3.27	41.36	84
	Posttest	13.61	64.00	84
Implicit	Pretest	−1.84	44.90	120
	Posttest	9.44	47.15	120

Figure 17

Plot of F0 differences for sentence stress in two groups at the pre- and posttest
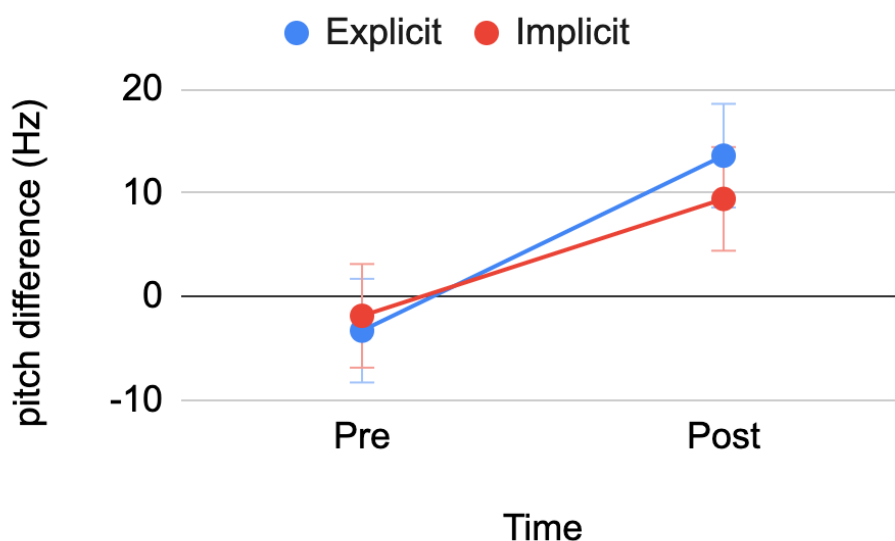
The two-way repeated measures ANOVA test results for F0 displayed a similar pattern of significance as found in intensity. Only the time factor showed significance, $p < .001$. Neither the group effect nor the time*group interaction displayed any significant results, $p = 0.441$ for time* group factor and $p = 0.821$ for group factor, as indicated in Table 30 and Table 31.

Table 30

ANOVA test results of F0 average for sentence stress – within subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Time	19596.999	1	19596.999	15.074	< .001
Time *Group	775.075	1	775.075	0.596	0.441
Residuals	262604.932	202	1300.024		

Note. Type III Sum of Squares

Table 31

ANOVA test results of F0 average for sentence stress – between subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Group	184.521	1	184.521	3582.881	0.821
Residuals	723742.032	202	3582.881		

Note. Type III Sum of Squares

Yet, seen from the statistics of the post-hoc tests in Table 32, the significant difference was only obtained in the explicit group between the pre and post ($t = -3.034$, $p < .05$), whereas the implicit group did not show a significant improvement ($t = -2.424$, $p = 0.081$). Same as for intensity values, neither the pretest or the posttest exhibited significant differences between the two groups.

Table 32

Post-hoc results of F0 for sentence stress in two groups at the pre- and posttest

		Mean Difference	SE	t	Pholm
Explicit, Pre	Implicit, Pre	−1.434	7.029	−0.204	1.000
	Explicit, Post	−16.883	5.564	−3.034	0.016
	Implicit, Post	−12.716	7.029	−1.809	0.214
Implicit, Pre	Explicit, Post	−15.448	7.029	−2.198	0.115
	Implicit, Post	−11.281	4.655	−2.424	0.081
Explicit, Post	Implicit, Post	4.167	7.029	0.593	1.000

Note. P-value adjusted for comparing a family of 6

4.2.3. Analysis of duration for contrastive stress

Concerning the duration analysis of contrastive sentence stress, the percentage values were obtained by using the duration of the contrastive word to divide that of the other words within the same sentence. The words used are the same as those for the

intensity and peak analysis. Table 33 shows the means of percentage generated by all subjects in each sentence in the pretest and posttest across the two groups.

Similar to word stress, the duration ratio values for sentence stress in most cases decreased in the posttest compared with the pretest, except a few cases marked in red. This means that participants mostly either pronounced the target tonic stress with longer duration or shortened the unstressed word. Just as in the figures for duration ratios of word stress, many percentage figures in Table 19 are also larger than 100%, which indicates that participants might have improperly produced the unstressed words with longer duration compared to the tonic stress words.

Table 33

Means of duration ratios for sentence stress in two groups at the pre- and posttest

Words for division	Explicit group		Implicit group	
	Percentage in Pre	Percentage in Post	Percentage in Pre	Percentage in Post
S1				
he / doctor	43%	42%	42%	45%
I / nurse	43%	43%	40%	48%
S2				
badminton / playing	153%	138%	149%	139%
badminton / watching	196%	200%	183%	174%
S3				
here / traveling	92%	74%	63%	58%
here / living	105%	88%	82%	78%
S4				
prefer / car	140%	112%	134%	108%
prefer / bike	101%	98%	123%	116%
S5				
cleaned / he	179%	140%	209%	184%
not / me	66%	69%	86%	72%
S6				
morning / cold	102%	99%	119%	120%
afternoon / hot	202%	160%	217%	189%

Table 34 presents the statistics of duration ratios between the two test points in terms of means and standard deviations. Both groups showed a decrease in the ratios in the posttest compared with the pretest; in the explicit group: *M* = 118.6%, *SD* = 65.6 in the pretest and *M* = 105.4%, *SD* = 52.4 in the posttest, whereas the implicit group, on the other hand, obtained M =120.7%, SD = 67.2 in the pretest, and M = 111.0%, SD = 57.3 in the posttest. The decrease between the pre- and posttests for the explicit group is 13.2% while in the implicit group it is 9.7%. Such a drop in percentage values means that either the stressed words was pronunced with longer duration or the unstressed word was shorter in the posttest.

Table 34

Descriptive statistics of duration ratios for sentence stress in two groups at the pre- and posttest

		M	*SD*	*N*
Explicit	Pretest	118.6%	65.6	84
	Posttest	105.4%	52.4	84
Implicit	Pretest	120.7%	67.2	120
	Posttest	111.0%	57.3	120

Figure 18 illustrates the changes in the means of duration ratios from the pretest to the posttest in each group. Both groups exhibited a downward trend and there was no interaction between the groups.

Figure 18

Plot of duration ratio means for sentence stress in two groups at the pre- and posttest
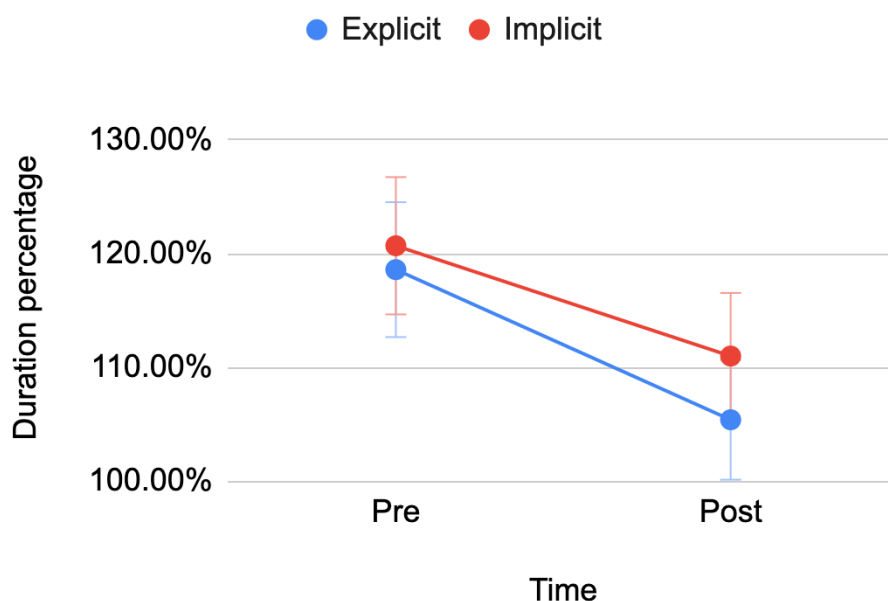
The two-way repeated measures ANOVA test confirmed the significance attributed to the time variable ($p < .001$). However, significance was not observed in either the time*group interaction ($p = 0.541$) or the group variable ($p = 0.639$). Detailed statistics for within-subject effects and between-subject effects are presented in Tables 35 and Table 36, respectively.

Table 35

ANOVA test results of duration ratios for sentence stress - within subjects effects

Cases	Sum of Squares	df	Mean Square	F	p
Time	1.292	1	1.292	15.255	< .001
Time *Group	0.032	1	0.032	0.375	0.541
Residuals	17.108	202	0.085		

Note. Type III Sum of Squares

Table 36

ANOVA test results of duration ratios for sentence stress - between subjects effects

Cases	Sum of Squares	*df*	Mean Square	*F*	*p*
Group	0.147	1	0.147	0.221	0.639
Residuals	134.254	202	0.665		

Note. Type III Sum of Squares

The results of the post-hoc tests, shown in Table 37, do not support the existence of a significant difference between the two groups in either the pre (p = 1.0) or the posttest (p = 1.0). Concerning the significance of changes within the groups, it is noteworthy that only the explicit group exhibited a significant difference from the pre- to the posttest (p < .05). While the implicit group did not quite reach statistical significance, it was notably close to it (p = 0.055).

Table 37

Post-hoc test results of the duration ratios for sentence stress in the two groups at the pre- and posttest

		Mean Difference	*SE*	*t*	Pholm
Explicit, Pre	Implicit, Pre	–2.1%	8.7%	–23.7%	1.000
	Explicit, Post	13.2%	4.5%	294.5%	0.022
	Implicit, Post	7.6%	8.7%	87.0%	1.000
Implicit, Pre	Explicit, Post	15.3%	8.7%	175.6%	0.321
	Implicit, Post	9.6%	3.8%	256.6%	0.055
Explicit, Post	Implicit, Post	–5.6%	8.7%	–64.9%	1.000

Note. P-value adjusted for comparing a family of 6

4.3. Results of production data rated by native judges

This section presents the results of participants' production data rated by 12 native English speaker judges. Firstly, the results of the reliability analysis are shown. Subsequently, the descriptive statistics presents the mean scores of both groups in comprehensibility and accentedness. Finally, two-way repeated measures ANOVA tests were performed to explore the significance of the differences across groups and between pre- and posttests for comprehensibility and accentedness, respectively. In instances where significance was observed in the AVONA tests, post-hoc tests were conducted for paired comparisons.

4.3.1. Results of the intraclass correlation coefficient

Table 38 presents the ICC values with 95% confident interval in each group for each test, specifically regarding comprehensibility and accentedness. A total of 8 tests were conducted. Notably, all values exceed 0.80 and many even surpass 0.90, which indicate a high level of reliability[7].

Table 38

Intraclass correlation reliability values for each test measured

Group	Rating aspect	Test	Reliability ICC values	95% confidence interval
Explicit group	comprehensibility	pretest	0.903	0.747-0.980
		posttest	0.881	0.690-0.976
	accentedness	pretest	0.942	0.850-0.988
		posttest	0.834	0.567-0.966

---

[7] Based on the 95% confident interval of the ICC estimate, values less than 0.5, between 0.5 and 0.75, between 0.75 and 0.9, and greater than 0.90 are indicative of poor, moderate, good, and excellent reliability, respectively.

Implicit group	comprehensibility	pretest	0.940	0.865-0.982
		posttest	0.918	0.815-0.976
	accentedness	pretest	0.921	0.823-0.977
		posttest	0.938	0.861-0.982

4.3.2. Results of descriptive statistics, ANOVAs and post-hoc tests

In order to test the significance of the difference for the group and time variables, two two-way repeated measures ANOVA tests were conducted for comprehensibility and accentedness, respectively. There were 7 productions from the explicit group and 10 from the implicit group, rated by 12 judges on a scale of 7 points, so in total the explicit group obtained 84 valid scores, and the implicit group 120 scores. Table 39 presents the means of the comprehensibility scores for each group between the two time points and the corresponding standard deviations. Overall, the implicit group performed better than the explicit group either at the pre- or posttest. However, when observing the changes within each group, the explicit group has made some improvements in comprehensibility from the pre- to the posttest (pre $M = 4.31$, $SD = 1.56$ vs. post $M = 4.49$, SD = 1.47), but in the implicit group, a slight decrease in scores is seen (pre $M = 4.89$, $SD = 1.47$ vs. post $M = 4.85$, $SD = 1.60$). Figure 19 presents the plot of mean scores between the pre and post tests across the two groups.

Table 39

Descriptive statistics for comprehensibility scores in two groups at pre- and posttest

		M	*SD*	*N*
Explicit	Pre	4.31	1.56	84
	Post	4.49	1.47	84
Implicit	Pre	4.89	1.47	120
	Post	4.85	1.60	120

Figure 19

Plot of the ANOVA test for comprehensibility scores
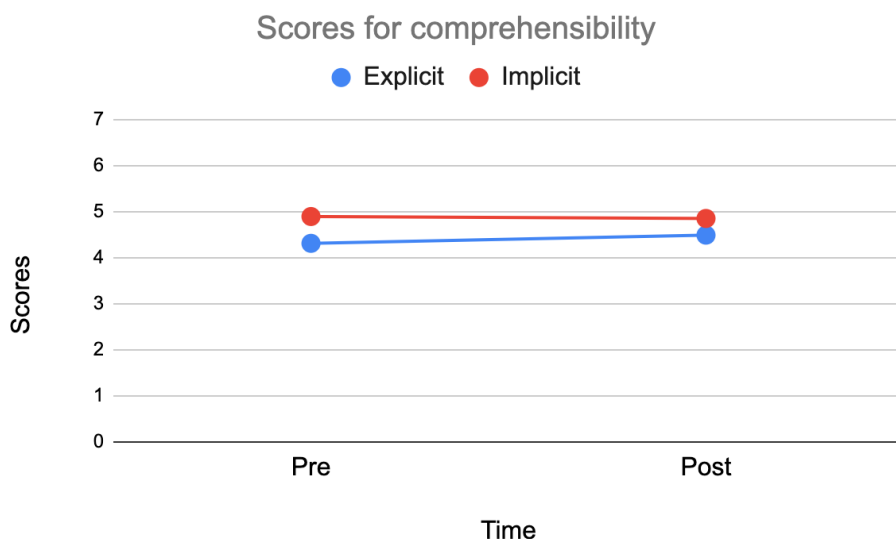
A two-way repeated measures ANOVA was run to explore whether any significance was present. The two independent variables were time as the within-subjects factor and group as the between-subjects factor. The means of the raters' comprehensibility scores were the dependent variable. As shown in Table 40, the result was not significant for time, $F= 0.634$, $p = 0.427$, but it was significant for group, $F = 5.396$, $p = 0.021$.

Table 40

ANOVA test results for the means of comprehensibility scores

	Effect	df	Mean square	F	p
Within-subject	Time	1	0.463	0.634	0.427
Between-subject	Group	1	22.019	5.396	0.021

In order to ascertain where significance occurred between the two groups, a Bonferroni post-hoc tests were calculated for multiple comparisons. Results are shown in Table 41. The statistics show there was no significant difference between the two

groups in either the pre- or posttest. Most importantly, the changes in each group were not significant either. This appears contradictory to the significant effect of the time variable in the ANOVA result, but it might be caused by the approximate values for significance in the pretest and posttest. As depicted in Figure 19, the implicit group seemed to have entered the experiment with a higher level of comprehensibility. However, such variations were unpredictable and beyond experimental control, as participants were recruited based on their overall level of English and allocated to different groups randomly. The focal point of attention should be directed toward the improvements observed within each group where distinct approaches were applied, as the primary objective of the investigation was the effectiveness of treatments within each group.

Table 41

Post-hoc test for comprehensibility across the two groups and between the two tests

		Mean Difference	*SE*	*t*	Pholm
Explicit, Pre	Implicit, Pre	-0.582	0.221	-2.638	0.053
	Explicit, Post	-0.179	0.132	-1.354	0.354
	Implicit, Post	-0.540	0.221	-2.450	0.075
Implicit, Pre	Explicit, Post	0.404	0.221	1.829	0.274
	Implicit, Post	0.042	0.110	0.378	0.706
Explicit, Post	Implicit, Post	-0.362	0.221	-1.640	0.306

As far as accentedness is concerned, Table 42 presents the means of the obtained scores for each group between the two tests and the corresponding standard deviations. As observed, both groups have made some small improvements in accentedness from the pre- to the posttest, in the explicit group, pre $M = 2.79$, $SD = 1.36$ vs. post $M = 2.87$, SD = 1.32); and in the implicit group, pre $M = 3.22$, $SD = 1.42$ vs. post $M = 3.27$, $SD = 1.46$). Figure 20 presents the plot of mean scores between the pre- and posttests across the two groups. Like the plot pattern of comprehensibility, accentedness in the implicit

group was rated slightly higher than the explicit group at the pre- and posttest, but within each group, there was not a notable change in scores.

Table 42

Descriptive statistics for accentedness scores in two groups at pre- and posttest

		M	*SD*	*N*
Explicit	Pre	2.79	1.36	84
	Post	2.87	1.32	84
Implicit	Pre	3.22	1.42	120
	Post	3.27	1.46	120

Figure 20

Plot of the ANOVA test for accentedness scores
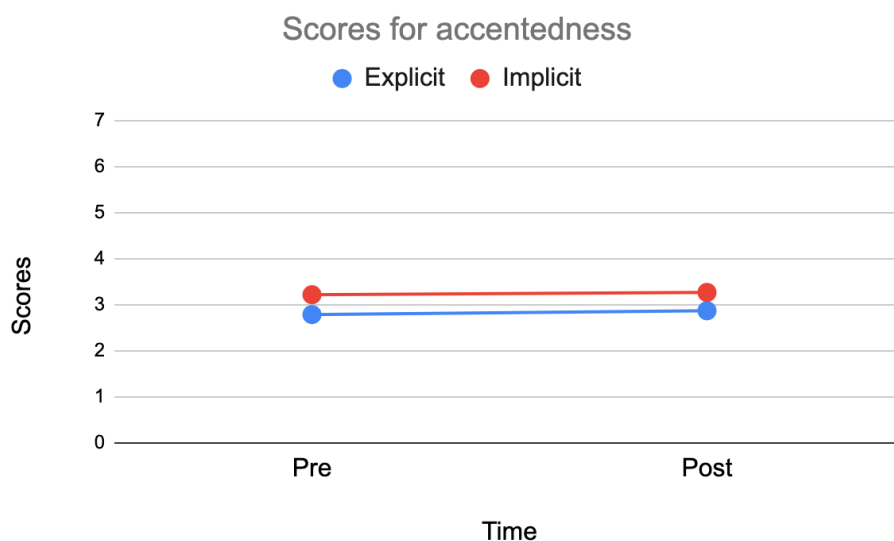
The second two-way repeated measures ANOVA test was conducted for the accentedness scores across groups from the pre- to the posttest. As in the previous tests, time was the within-subjects factor and group was the between-subjects factor. Yet, the dependent variable of this test was the means of the raters' accentedness scores. As

shown in Table 43, the result was not significant for the time factor, $F = 0.647$, $p = 0.422$, but for the group factor, there was a significant difference, $F = 5.219$, $p = 0.023$.

Table 43

ANOVA test results for the means of accentedness scores

	Effect	*df*	Mean square	*F*	*p*
Within-subject	Time	1	0.439	0.647	0.422
Between-subject	Group	1	16.961	5.219	0.023

In the same vein, a post-hoc test was subsequently carried out for multiple comparisons, and the results are presented in Table 44. Similar to the results for comprehensibility, there was no significant difference between the two groups in either the pre- or posttest, nor did significance occur within groups from the pre to the post test. The results seemed to contradict the significance of the group variable in the ANOVA result, which might relate to the fact that the implicit group was consistently better than the explicit group. Nevertheless, as shown in Figure 19, the difference between groups was minimal. What should draw more attention is the change from the pre- to the posttest within each group.

Table 44

Post-hoc test for accentedness across two groups and between two tests

		Mean Difference	*SE*	*t*	Pholm
Explicit, Pre	Implicit, Pre	-0.431	0.199	-2.638	0.157
	Explicit, Post	-0.083	0.127	-1.354	1.000
	Implicit, Post	-0.481	0.199	-2.450	0.099
Implicit, Pre	Explicit, Post	0.348	0.199	1.829	0.247
	Implicit, Post	-0.050	0.106	0.378	1.000
Explicit, Post	Implicit, Post	-0.398	0.199	-1.640	0.188

4.4. Results of the qualitative questionnaire

In order to make students reflect on the pronunciation training they received, participants were asked to answer a questionnaire upon the completion of the training course. Both 5-point Likert scale questions and open-ended questions were included to gather students' opinions on the importance of English pronunciation and their feedback on the effectiveness of the course. Apart from some identical general questions applied to both groups, there were also a series of specific questions tailored exclusively for each experimental group to collect their feelings towards receiving the treatment. In the subsequent sections, results for 5-point Likert scale general questions answered by both groups will be selectively presented in tables first. In order to compare responses from the two groups in a neat and logical manner, the order of some questions is slightly changed. For the same questions answered by both groups, t-tests were used to compare the group difference in responses. Besides, a selection of group-specific results will be presented along with some open-question comments related to participants' feedback on the course. Each group had 11 participants who answered the questionnaire. All participants' responses were considered and will be displayed in this section.

4.4.1. General results of participants' views on pronunciation

Table 45 summarizes the results obtained from both groups correlated with the participants' attitudes towards the importance of English pronunciation and the participants' views on their own pronunciation.

Table 45

Participants' feedback from both groups to pronunciation-related questions

Questions	Explicit group		Implicit group	
Q1: How important do you think pronunciation is for English learning?	1-*Not important at all*	0%	1-*Not important at all*	0%
	2	0%	2	0%
	3	9.09%	3	18.18%
	4	9.09%	4	18.18%
	5-*Very important*	81.82%	5-*Very important*	63.64%
Q2: What did you think of your pronunciation before our course?	1-*Very bad*	9.09%	1-Very bad	18.18%
	2	18.18%	2	18.18%
	3	54.55%	3	54.55%
	4	9.09%	4	9.09%
	5-*Very good*	9.09%	5-Very good	0%
Q3: To what extent do you think accurate pronunciation is necessary to communicate with native speakers?	1-*Not necessary at all*	0%	1-*Not necessary at all*	0%
	2	0%	2	0%
	3	9.09%	3	27.27%
	4	54.55%	4	45.45%
	5-*Very necessary*	36.36%	5-*Very necessary*	27.27%

It is evident that the importance of pronunciation for English learning was acknowledged by all participants, 81.82% from the explicit group and 63.64% from the implicit group regarded pronunciation as *very important*, and no one claimed it was *not important*. To investigate if there was a significant difference between the two groups, an independent-samples t-test was performed. The mean score of the explicit group was 4.73 and the mean of the implicit group was 4.46, as shown in Table 46. No significant difference was found between the two groups' responses ($t = 0.866$, $p = 0.397$), as indicated in Table 47.

Table 46

Mean scores and standard deviations for Question 1 between the two groups

	Group	N	Mean	SD	SE	Coefficient of variation
Score	Explicit	11	4.73	0.65	0.20	0.14
	Implicit	11	4.46	0.82	0.25	0.18

Table 47

Independent sample t-test of mean scores for Question 1 between the two groups

	t	df	p
Score	0.866	20	0.397

In terms of participants' self-perception of their pronunciation before entering the training course, most of them held a negative or neutral attitude toward their pronunciation in both groups. Only 18.18% of the participants from the explicit group rated their pronunciation at 4 points or above, and 9.09% did so in the implicit group. Nobody believed their pronunciation to be *very good*.

An independent-samples t-test was performed to test if there was a significant difference in their self-perception between the two groups. The mean score of the explicit group was 3.0 and the mean of the implicit group was 2.55, as shown in Table 48. No significant difference was found between the two groups' responses ($t = 1.00$, $p = 0.329$), as indicated in Table 49.

Table 48

Mean scores and standard deviations for Question 1 between the two groups

	Group	N	Mean	SD	SE	Coefficient of variation
Score	Explicit	11	3.00	1.18	0.36	0.39
	Implicit	11	2.55	0.93	0.28	0.37

Table 49

Independent sample t-test of mean scores for Question 2 between the two groups

	t	df	p
Score	1.000	20	0.329

However, the majority of participants recognized the necessity of accurate pronunciation for effective communication with native speakers. In the explicit group, a substantial 90.91% of participants rated pronunciation as crucial with a score of 4 points or higher, while 72.72% of those in the implicit group shared the same view. Importantly, nobody in either group deemed pronunciation to be unnecessary.

The mean score of the explicit group was 4.27 and the mean of the implicit group was 4.00, as shown in Table 50. The t-test results in Table 51 indicate that there was no significant difference between the two groups' responses ($t = 0.896$, $p = 0.381$).

Table 50

Mean scores and standard deviations for Question 3 between the two groups

	Group	N	Mean	SD	SE	Coefficient of variation
Score	Explicit	11	4.27	0.65	0.20	0.15
	Implicit	11	4.00	0.78	0.23	0.19

Table 51

Independent sample t-test of mean scores for Question 3 between the two groups

	t	df	p
Score	0.896	20	0.381

4.4.2. Group-specific results of participants' views on the course

The following questions presented in Table 52 are related to participants' views about the training course, specifically, how easy they feel the course is and how useful it is for improving their pronunciation.

Table 52

Participants' views from both groups about the training course

Questions	Explicit group		Implicit group	
Q4: How easy do you feel it is to follow our course?	1-*Very difficult*	0%	1-*Very difficult*	0%
	2	0%	2	0%
	3	45.45%	3	9.09%
	4	27.27%	4	63.64%
	5-*Very easy*	27.27%	5-*Very easy*	27.27%
Q5: How useful do you think our training course is for improving your pronunciation?	1-*Not useful at all*	0%	1-*Not useful at all*	0%
	2	0%	2	0%
	3	0%	3	18.18%
	4	36.36%	4	63.64%
	5-*Very useful*	63.64%	5-*Very useful*	18.18%

When assessing the ease in following the course, it can be seen that about half of the population in the explicit group found the course *easy* (27.27%) or *very easy* (27.27%), while the remaining 45.45% expressed a neutral stance on its difficulty. No one in the group deemed the course to be difficult. Similarly, nobody in the implicit group found the imitation practice to be demanding. In comparison to the explicit group, an even larger percentage in the implicit group regarded the class activities as *easy* (63.64%) or *very easy* (27.27%), with only 9.09% rating it as moderately difficult with a score of 3 points.

The t-test results reveal that the mean score of the explicit group was 3.82 and the mean of the implicit group was 4.00, as shown in Table 53. There was no significant difference between the two groups regarding whether the course was easy to follow ($t = -0.559$, $p = 0.582$), as shown in Table 54.

Table 53

Mean scores and standard deviations for Question 4 between the two groups

	Group	*N*	Mean	*SD*	*SE*	Coefficient of variation
Score	Explicit	11	3.82	0.87	0.26	0.23
	Implicit	11	4.00	0.63	0.19	0.16

Table 54

Independent sample t-test of mean scores for Question 4 between the two groups

	t	*df*	*p*
Score	-0.559	20	0.582

Regarding participants' evaluation of the usefulness of the training, all participants in the explicit group felt satisfied with the training, with 36.36% considering it *useful* and 63.64% deeming it *very useful* for improving their pronunciation. In the implicit group, while the course was not rated as highly as in the explicit group, a notable 63.64% found it *useful*, 18.18% deemed it *very useful*, and the remaining 18.18% held a neutral stance regarding its usefulness in improving pronunciation.

The mean score of the explicit group was 4.64 and the mean of the implicit group was 4.00, as shown in Table 55. The t-test revealed a significant difference between the two groups' responses ($t = 2.609$, $p = 0.017$), as indicated in Table 56.

Table 55

Mean scores and standard deviations for Question 5 between the two groups

	Group	N	Mean	*SD*	*SE*	Coefficient of variation
Score	Explicit	11	4.64	0.51	0.15	0.11
	Implicit	11	4.00	0.63	0.19	0.16

Table 56

Independent sample t-test of mean scores for Question 5 between the two groups

	t	df	p
Score	2.609	20	0.017

In order to find out which specific aspect posed challenges for the participants, an open-ended question was posed: "If it is not easy for you, can you name the difficulties you have encountered?" Some representative responses from each group are quoted as presented in Table 57.

Table 57

Select responses from both groups on the difficulties that participants encountered

Explicit group	Implicit group
S5: "Sometimes I can't speak fluently. Some (of my) accents are incorrect. (I am) Not familiar with some fundamental pronunciation skills."	S4: "I can't answer which words should be stressed."
S9: "I know the rules, but I cannot pronounce the words or phrases well."	S8: "I lack some specialized knowledge."
S11: "I used to follow my feel to rise tone or descend one, before I study this course."	S11: "I am not aware where the stress should be."

4.4.2.1. Questions addressed to the explicit group exclusively

Each group also answered some tailored questions related to the distinctive teaching approach and practice used in training. For example, in the explicit group, the following question was asked: "How much can you understand the pronunciation rules taught in the course?" The options ranged from 1 = *none of them* to 5 = *all of them*. The vast majority of participants reported the rules were quite understandable, in total 90.90% of participants rated them at 4 points and above, as shown in Figure 21.

Figure 21

Explicit group's responses to "How much can you understand the pronunciation rules taught in the course?"
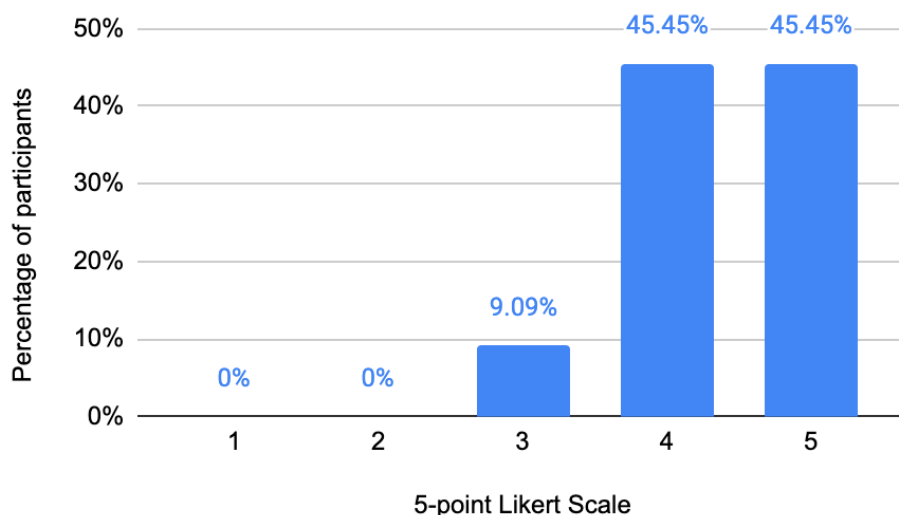
Afterwards, the participants were asked if they were willing to continue improving other aspects of their English pronunciation in this way. All of the participants expressed either *willing* (36.36%) or *very willing* (63.64%). The responses are shown in Figure 22.

Figure 22

Explicit group's responses to "Are you willing to continue improving other aspects of your English pronunciation in this way?"
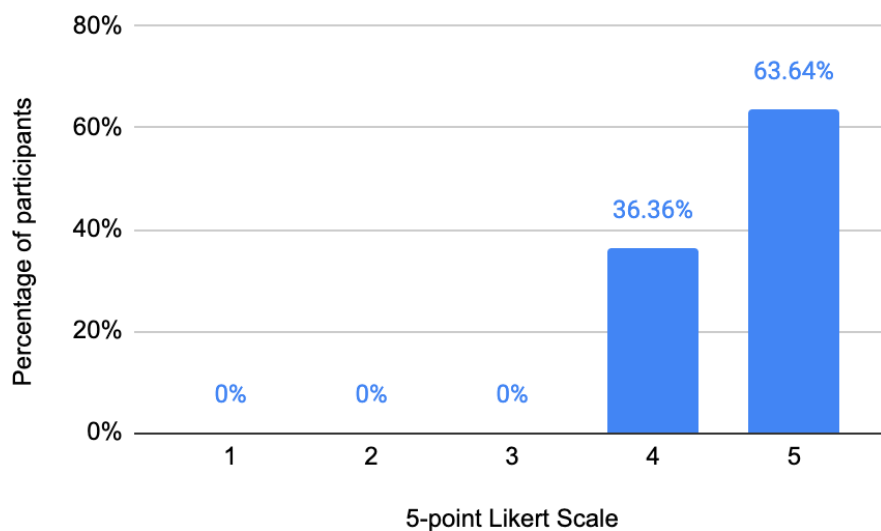
The participants in the explicit group were also questioned if they would like to have native speakers' recordings to imitate as a supplement for self-practice. It seems most of them had an open mind about this technique for improving pronunciation. The options ranged from 1 = *not willing* to 5 = *very willing*. 63.64% chose *very willing* and 18.18% *willing* as shown in Figure 23.

Figure 23

Explicit group's responses to "Would you like native speakers' recordings to imitate for self-practice?"
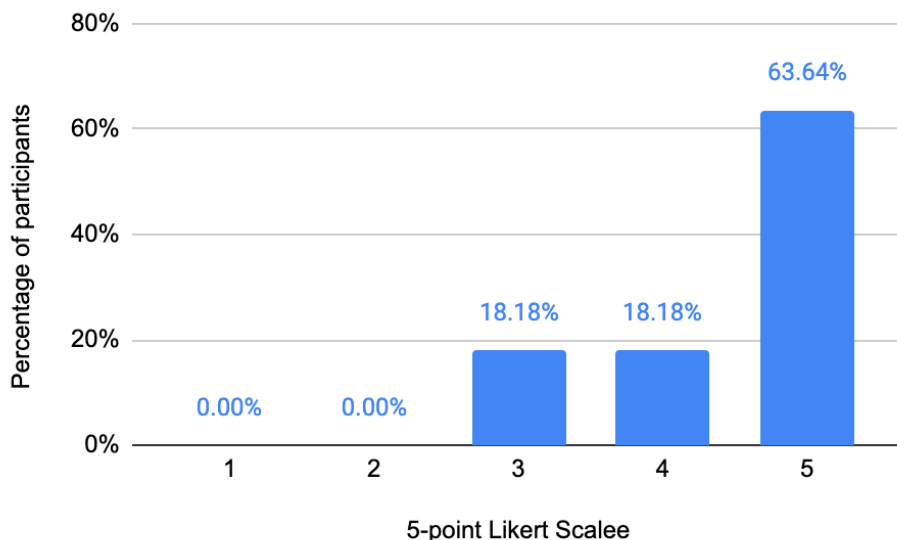
Subsequently, an open-ended question was asked to find out how the course benefitted the participants from their perspective: "In which way do you think this course has helped you?" Participants' representative responses were selected and are presented below.

> S1: "I get more pronunciation's theoretical knowledge."
> S2: "It helped me in pronunciation and sentence breaking in spoken English, which benefit me a lot."
> S3: "This course discussed different types of stress more specifically, enabling me to know more detailed pronunciation rules."
> S5: "Pay much more attention to pronunciation when speaking. Learn and try to stress sentences."
> S10: "The tips learnt from the course is really useful and can be used in lots of situations."

Finally, participants' additional suggestions for the course were collected by an open-ended question: "Do you have any suggestion to help us improve the course?" A few participants' responses are quoted below. It appears many participants felt the course delivered practical knowledge but was short of communication practice where they could apply the use of rules to speech:

S2: "I think the teacher has taught well, if (she) can add some everyday words, which (it) will (be) better."

S3: "Involving more practice may be helpful to students mastering the pronunciation rules."

S4: "Add more exercises and speaking activities."

S5: "Much more practice and longer class time."

S6: "Give students more opportunity to speak English with others. Practice makes perfect. We need an English-speaking environment."

4.4.2.2. Questions addressed to the implicit group exclusively

The following questions were addressed to the implicit group exclusively. Imitating native recordings was the major methodology used in this group, so it is essential to take into account how much participants engaged in the practice in order to interpret the results. Figure 24 presents participants' responses to the question "How much have you followed the audios and done the imitation practice?" The options ranged from 1 = *0%* to 5 = *100%*. Slightly more than half of the population claimed to have done 100% imitation practice with the provided native recordings and two participants failed to do more than 50% imitation practice, as illustrated in Figure 24.

Figure 24

Implicit group's responses to "How much have you followed the audios and done the imitation practice?"
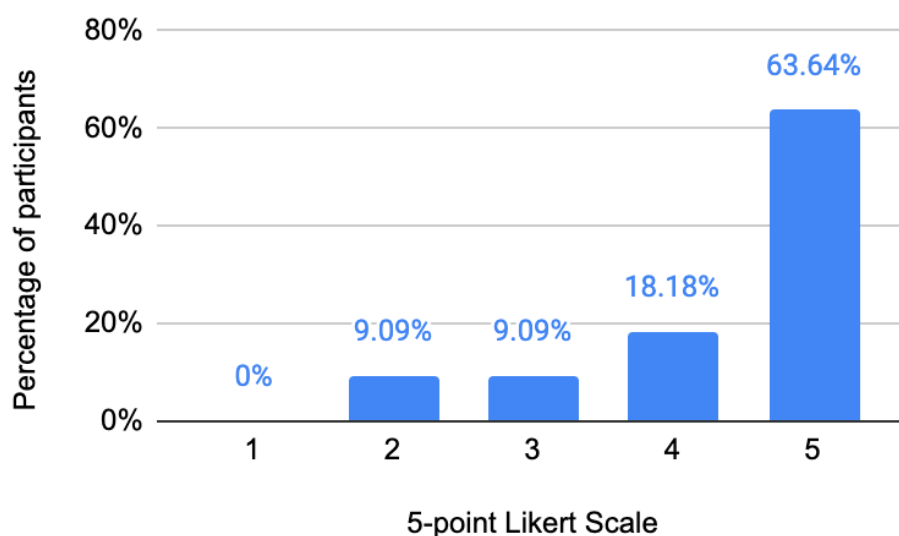
Regarding participants' views on the imitation practice, the question "To what extent do you think this imitation practice has improved your pronunciation?" was addressed. The options ranged from 1 = *not at all* to 5 = *very much*. Slightly more than half of the participants were positive about the effect of the imitation practice, including 18.18% chose *very much* and 36.36% *much*. The others felt neutral about this method of training for improving pronunciation, accounting for 45.45%. Nobody regarded it as useless. Figure 25 illustrates the distribution.

Figure 25
Implicit group's responses to "To what extent do you think this imitation practice has improved your pronunciation?"
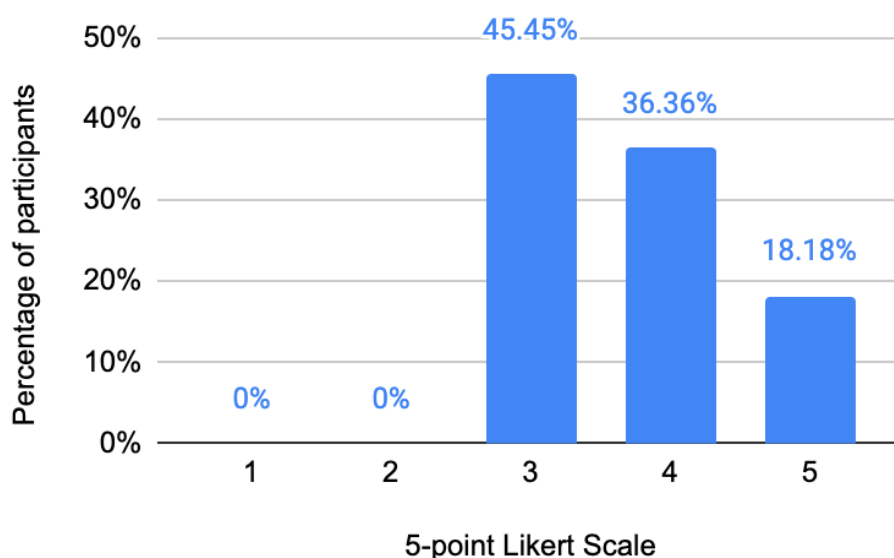
Just like in the explicit group, in order to gain a deeper understanding of the ways in which the implicit approach was beneficial to the participants, an open-ended question was posed: "In what manner do you believe this course has aided you?" Representative responses are presented selectively as follows.

> S1: "Improve my pronunciation. I can know more knowledge about pronunciation, such as what word should be stressed."
> S3: "This course improved my intonation, for example, which word should be emphasize(d)".
> S4: "To stress the words to express our emotion. Be more confident to communicate in English. Communication should (have) pause."
> S5: "To notice that the sentences that we use in our daily lives also have rhyme."

S8: "Emphasize some words when reading."

S11: "Practiced my listening and speaking." (translated from Chinese transcript)

Finally, in terms of further suggestions towards the course, only a few participants left some feedback, quoted as follows:

S2: "Should invite more students to answer questions as much as possible to make class livelier."

S4: "Sometimes I had problems understanding the content. With explanations in Chinese would be better." (translated from Chinese transcript)

S6: "Hope the course lasted longer and covered more aspects of English learning."

CHAPTER V
DISCUSSION AND CONCLUSIONS

The aim of the present study was to compare two commonly used methods of teaching pronunciation (implicit versus explicit instruction) to Chinese adult EFL learners. The objective was to determine which one of the two is more effective in teaching English word stress and contrastive sentence stress. This chapter firstly examines and interprets the results presented in Chapter IV with respect to the proposed research questions, and subsequently provides a conclusion of the main research findings, as well as listing limitations of this study, alongside with the implications for the research of the field.

This chapter is divided in five sections. Section 5.1 summarizes and discusses the results of the acoustic speech analysis of the words and sentences reading. Section 5.2 examines the results of the judges' evaluations of the spontaneous talk produced by participants. Section 5.3 briefly discusses the participants' responses from the satisfaction surveys. Finally, Section 5.4 presents an overview of the major findings of this study and Section 5.5 addresses the limitations of the study and meanwhile propose some suggestions for future research.

5.1. Discussion of the acoustic analysis results

5.1.1. Discussion on RQ1

In answer to the first research question, the graphs depicting changes in intensity, F0, and duration cues, presented in Chapter IV, have revealed positive improvements from the pre-test stage to the post in both word stress and sentence stress scenarios. Specifically, intensity and F0 difference values exhibited an upward trend, while the duration ratio decreased. These findings indicate an increased awareness among participants to place more prominence on the target tonic syllables/words with greater intensity, higher F0 and longer duration. Despite the fact that statistical significance

101

was not consistently observed in all cases, it appears that the prosody training effectively impacted participants in both groups.

In particular, the explicit group displayed a significant improvement in the posttest for contrastive sentence stress, as shown by the statistical analysis of all three acoustic cues—intensity, F0, and duration. However, it is worth noting that the duration effect did not reach statistical significance for word stress in this group. In contrast, the differences observed in the implicit group were not consistently significant for both word stress and sentence stress. For instance, regarding word stress in the implicit group, only the F0 cue exhibited a significant difference, with the intensity and duration cues failing to reach statistical significance. In terms of sentence stress, significance was only found in intensity values with no statistical significance observed for either F0 or duration. Table 58 provides a summary of the statistical significance results for each acoustic cue from the pre- to the posttest, differentiating between word stress and sentence stress, respectively. Ticks represent significance, whereas crosses indicate non-significance.

Table 58

Statistical significance results of each acoustic cue for word stress and contrastive sentence stress

Word stress			
	Intensity	F0	Duration
Explicit Group	✅	✅	❌
Implicit Group	❌	✅	❌
Sentence stress			
Explicit Group	✅	✅	✅
Implicit Group	✅	❌	❌

The variability in the distribution of significance can be interpreted as a complementary observation to the ongoing debate and lack of consensus within the literature concerning the hierarchy of acoustic correlates for English lexical stress. Despite the customary belief that stressed syllables are usually produced with relatively higher F0, greater intensity, and longer duration compared to unstressed syllables, as presented in Section 2.6, the inconsistencies observed in acoustic research are not uncommon (e.g. see Beckman, 1986; Gordon & Roettger, 2017; Jung & Rhee, 2018; Ladd, 2008; Roach, 2009). Most researchers agree that the primary cue for stress in English in both natural and synthesized speech is relative pitch prominence, followed by length and amplitude (Roach, 2009). However, factors like noise and pitch accent could also lead to misrepresentation of intensity and F0 cues. As predicted by Ladd (2008), investigating F0 to define stress can be challenging due to the co-occurrence of word-level prominence and post-lexical tonal events. Additionally, since the data for this study were collected remotely, the sound quality of the recordings was not always ideal. It could be possible that external variables, such as noise and other environmental factors influenced the intensity and pitch values slightly.

Regarding duration measurements, it was observed that there were decreases in duration ratios, specifically, the duration of the unstressed syllable/word divided by that of the primary stressed syllable/word for word stress and contrastive sentence stress. However, the results of duration changes from the pre to the posttest, if seen in isolation from intensity and F0, may not serve as a very powerful indicator to determine definite improvement. This is because most percentage figures exceed 100%, which appears to indicate that participants were not producing the primary stress with a longer duration compared to the relatively less stressed ones within the same word or sentence.

Seeing the unexpected duration results, it seems that Chinese speakers are not used to extending the duration when uttering stressed syllables. This phononemon could be attributed to several factors. First of all, as pointed out by Bu & Zhou (2021), Chinese learners often face considerable challenges with English stress assignment. These difficulties arise from the intricate and irregular nature of English stress placement,

whether at the word or sentence level, and also from the lack of similar stress patterns between the two languages.

As previously mentioned, Chinese is a syllable-timed language, where each syllable is typically pronounced with equal time duration, regardless of whether it is stressed or unstressed. This inherent characteristic of the native language can manifest itself in the pronunciation of a second or foreign language, particularly when such habits become fossilized in adult learners. Changing these habbits with a short-term training course has proven to be a challenging endeavor.

Additionally, it is important to bear another factor in mind, which is that the duration values in this study were derived from syllable segmentation for measuring word stress (instead of vowels) and from word segmentation for sentence stress on PRAAT. Since the same syllables and words were consistently selected in both the pretest and the posttest for paired-sample comparisons, this ensures that the accoustic changes in duration are not influenced by the segmentation decision.

However, under such circumstances, it is not suprising to observe instances where a stressed monosyllable was pronounced shorter than an unstressed disyllable. The number of speech sounds could account for why some duration percetages exceed 100%. Another influencing factor could be *final lengthening* (Ladd, 1996), namely, a prominent lengthening of the last sound. For example, in the word "analyze", the duration ratios indicate that the syllable "lyze" (Sy3) is two or even three times longer than the primary stress syllable "a" (Sy1).

Word	Explicit pre	Explicit post	Implicit pre	Implicit post
Analyze Sy3/Sy1	261.04%	287.94%	276.72%	329.12%

Note. The figures are extracted from Table 18.

This segmentation issue also extends to the sentence level, as exemplified in Sentence 6: "It is cold in the morning and hot in the afternoon." Here, the tonic stress words "cold" and "hot" contain fewer syllables compared to the other two words,

"morning" and "afternoon" in the division calculation. Consequently, this can result in the duration percentages exceeding 100%. However, as mentioned above, this does not afftect the results of the paired comparisons from the pre- to the posttest since the same words were consistently utilized for calculation.

Furthermore, it is crucial to emphasize that this segmentation matter primarily affects the interpretation of duration values, and has a minimal impact on the other two acoustic cues, namely intensity peak and F0 average, as both intensity and pitch are generally considered to show higher values in the nucleus syllable or word. An initial decision was made to segment syllables or words for analyzing intensity and F0. When calcaulating the duration ratios, for the sake of time-efficiency, we chose to utilize the pre-segmented syllables (for word stress) and words (for sentence stress), which were previously employed in intensity and F0 analysis.

Despite the contentious duration results, the ratios still show a decrease in the post-training phase compared to the pretest. In conjunction with the intensity and F0 results, this decrease can be viewed as evidence of improved pronunciation of stress.

5.1.2. Discussion on RQ2

As far as the second research question is concerned, despite some visible improvements in the implicit group, it appears the treatment employed in the explicit group yielded better results in enhancing participants' production of word stress and contrastive sentence stress, considering that the explicit group has reached more statistial significance acoustically. This finding aligns with existing literature that advocates for explicit pronunciation instructions (e.g. Khanbeiki & Abdolmanafi-Rokni, 2015; Koike, 2014; Naeini & Adni, 2017; Yakut, 2020). Notably, the results of the present study are consistent with the findings of Khanbeiki & Abdolmanafi-Rokni (2015), despite their focus being on consonants. With the help of explicit rules, learners are more likely to notice language features and build up a greater awareness to pronounce stressed syllables more effectively, leading to faster improvement in the target language. Many SLA theorists propose that the first step in learning a second or

foreign language is to intentionally expose learners to formal rules of the target language features, and make them consciously aware of them, because metalinguistic awareness is central to interlanguage development (Schmidt, 2001). Abdolmanafi-Rokni (2009) also concluded that explicit instruction can increase the learning speed and generate more positive effects on pronunciation in a classroom setting in a short run in comparison to implicit instruction.

The effectiveness of explicit phonetic instruction has been confirmed in many more studies, with respect to various languages. As summarized in Hamzah's review (2014), this approach has been shown to improve L2 production accuracy in a number of languages, including English, French, German, and Spanish, leading to the belief that most adult L2 learners do not achieve native-like pronunciation without the help of explicit phonetic instruction (Fullana, 2006). It has been indicated that explicit instruction is more effective and time-efficient than implicit instruction for L2 learners' pronunciation acquisition (Gordon et al., 2013; Khanbeiki & Rokni, 2015).

Although in Peltekov's study (2020) the explicit group tended to perform worse than the implicit group and, in some cases, was even outperformed by the control group, the author argued that this should not be misinterpreted as evidence of the ineffectiveness of explicit phonetic rules, as students in the explicit group were possibly overwhelmed by the number of rules they had to learn in a short period, which led to the discrepancy. Previous research also suggests that only simple and clear-cut rules can be learned explicitly, whereas more complex and abstract rules are still subject to implicit learning (DeKeyser, 2003). In the current study, the course content was streamlined and the rules were easily accessible to participants in the explicit group, which might have facilitated their acquisition, hence improving their performance in the posttest. Furthermore, it is worth noting that Peltekov's study focused on learners' improvement in accent and comprehensibility, measured through a pretest and posttest. Thus, his conclusions cannot entirely apply to our findings on acoustic stress production.

On the other hand, the implicit approach, despite its observable advantages, did not benefit learners to the same extent. This could be explained by connectionist

theories, which suggest that implicit knowledge requires exposure to massive amounts of input and acquisition is a slow, organic, elaborate process (Ellis, 2015). The present study provided only short-term training (i.e., 4 hours in total distributed over 4 weeks) and the native recordings used for imitation practice could not meet the criterion of "massive amounts of input". Therefore, it is logical to assume that the implicit group could not achieve as notable a progression as the explicit group.

Moreover, as Spaai & Hermes (1993) reported, if prosody is taught implicitly without clear explanations, i.e., simply by means of the "listen and repeat" method, it cannot really be learned. In the same way, Ellis (2009) suggested that the explicit knowledge acquired through explicit learning can assist the processes involved in implicit learning. Implicit instruction cannot be arbitrarily dismissed as ineffective, but under the circumstance of concurrent short-term training, explicit instruction emerges as the preferable choice, especially for older or adult learners.

Age should be another pivotal factor to consider when assessing the efficacy of implicit learning. According to DeKeyser (2003), adult learners perform much worse than children in implicit learning. Brown (2007) also stated that older learners are expected to improve less than younger counterparts, especially if pronunciation is taught implicitly. In contrast to L1 learners, the older L2 learners face the challenge of losing aural sensitivity and also have limited exposure to the prosodic systems of expressiveness appropriate to the target language. Kucukoglu (2012) observed that, after the critical period, learners grapple with the challenge of perceiving which syllables are stressed and which are unstressed. Rhythm and stress models of speech are predominantly acquired in childhood and arduous for an adult to change. The assumption of an implicit approach suiting younger learners rather than older ones share a certain correlation with Kissling's proposal (2013), who suggested that explicit pronunciation instruction might be more effective for advanced learners, whereas implicit instruction is more acceptable for beginners. Future research could involve children or beginner learners in a study comparing the efficacy of the two methods

under similar conditions and procedures. These results could provide valuable insights into designing pronunciation courses tailored to learners' ages or proficiency levels.

Apart from age and proficiency level, it is also important to consider the impact of other individual differences, such as learning style and motivation. According to Ellis (2009), implicit instruction does not guarantee implicit learning, and explicit instruction may not exclusively lead to explicit learning. It is possible that students with different learning styles follow their own inclinations, and their learning outcomes might not align entirely with the nature of the instruction they receive. For example, students who are more prone to learn implicitly might struggle to grasp a teacher's explicit explanations of specific rules but, if this explanation is provided in the L2, these explanations could serve as input for the implicit acquisition of other language features that learners subconsciously infer from the teacher's speech. By contrast, those accustomed to explicit instruction, a method commonly used in Chinese EFL classrooms, may find implicit learning confusing and fail to independently decipher the language rules.

Furthermore, motivation could exert an influence on the degree to which participants benefit from pronunciation instruction (Peltekov, 2020). Gardner and Lambert (1972) coined the concepts of integrative motivation (i.e., language learning for personal growth and cultural enrichment) and instrumental motivation (i.e., language learning for immediate and practical goals). Both types of motivations are believed to relate to success of L2 learning. Although motivation was not the primary focus of this study, through participant observation of class performance, it was witnessed that highly motivated participants displayed more involvement in the training course and accordingly generated better productions in the posttest. Ultimately, Macdonald et al. (1994) found no significant differences in the effects of four different teaching techniques on L2 learners' pronunciation, but they observed that "the individual learner may serve as a more powerful variable than does the instructional setting in the acquisition of pronunciation" (p. 95).

All in all, as Thomson & Derwing (2015) concluded in their review of 75 studies on L2 pronunciation instruction, there are numerous factors contributing to divergent outcomes observed from different studies, including individual learner differences, goals and foci of instruction, type and duration of instructional input, and assessment procedures. It is important to note that no study is an exact replica of its predecessors, which implies that variations in results should be anticipated as a natural consequence. However, these factors could offer valuable insights for future researchers, guiding them in determining essential considerations when interpreting their findings.

Finally, taking into account the practicality and effectiveness of both teaching approaches, it is anticipated that a combination of these two methods might yield doubly positive effects. Roccamo (2015) explored this possibility and achieved significant positive outcomes. In that study, the experimental group, exposed to both implicit and explicit teaching methods, exhibited a notable improvement in comprehensibility compared to the control group, which received no pronunciation instruction. Following the design of this study, Peltekov (2020) revised the experiment and separated explicit instruction and implicit instruction so as to assess the two methods' individual effectiveness among beginner learners of Germans in terms of accent and comprehensibility. However, contrary to earlier expectations, the results show that neither the implicit nor the explicit experimental group in the experiment improved more than the control group. Peltekov concluded that the combined use of both instructional approaches proves more beneficial than employing either of them in isolation. Regrettably, as he pointed out, incorporating the two types of instruction is often unfeasible due to time limitations.

Likewise, Yang (2016) has also highlighted the benefits of combining these two methods for pronunciation training. In his study, despite the primary goal of investigating the efficacy of imitation on enhancing prosody, he tended to explicitly explain the transcripts to participants to mitigate task difficulty, prior to the activities of shadowing and mimicking authentic audios. Although it was not mentioned in detail what explicit instruction was provided, he concluded that the exposure to natural-

speech databases enables students to intuitively absorb the aspects of the language, which could enhance the explicit rules they have previously learned for the prosodic elements of the language. His study observed the significant positive effects of exposure to the suprasegmental levels of authentic speech on learners' fluency in production. On top of that, Yang pointed out the feasibility of utilizing this method for self-practice at home, preferably with the help of computer-assisted software like PRAAT, so that learners are able to self-evaluate their own production. Besides, this personalized practice accommodates each student's proficiency level and could significantly benefit highly motivated students particularly.

5.2. Discussion of perceived comprehensibility and accentedness

To answer RQ3: whether the training helped improve learners' perceived speech comprehensibility and accentedness, the descriptive statistics presented in Section 4.3 indicate that in terms of comprehensibility, although the implicit group started at a comparatively higher level, the explicit group improved slightly from the mean score of 4.31 to 4.49, whereas the implicit group dropped faintly from 4.89 to 4.85. Regarding accentedness, likewise, the implicit group entered the experiment with a slightly better accent, but both groups witnessed a raise in scores, the explicit group from 2.79 to 2.87, and the implicit group from 3.22 to 3.27. Thus, Hypothesis 3 appeared to be supported to some extent. However, no significance was detected between the pre- and posttest for either group; therefore, the effectiveness of a particular training approach for general comprehensibility and accentedness could not be seen as beneficial for participants from a statistical perspective.

Nevertheless, there are some justifiable reasons for the disparities observed between the acoustic analysis and human evaluations of comprehensibility and accentedness. Firstly, the intricate nature of determining factors for comprehensibility and accentedness in free speech suggests that overall improvements in learners' comprehensibility and accentedness may be difficult to achieve with short-term instruction. Despite clear acoustical differences, such changes may not be readily

discernible to the human ear, particularly in a manner significant enough to influence ratings of general comprehensibility and accentedness. Particularly, the degree of comprehensibility and accentedness comprises a blend of various attributes. Derwing & Munro (2015) emphasized that "segmental errors can also lead to misinterpretations of utterances and add to the perception of foreign accent" (p. 9). More recent studies (see Dahmen et al., 2023) have also indicated that there is a strong interplay between segmental and prosodic features when native listeners rate speech for intelligibility, comprehensibility, and degree of perceived foreign accentedness. Therefore, achieving increased comprehensibility and diminished foreign accent is unlikely to be solely realized through suprasegmental training.

Additionally, as pointed out by Kebboua Chaker (2019), "a task demand effect" and "test anxiety" are likely to interfere with students' performance in free production, as well as "rater fatigue effect" possibly affecting rating scores too. Hence, scores of comprehensibility and accentedness are influenced by multiple factors. However, the inability to achieve the significant improvements in these two constructs cannot be interpreted as failure of the pronunciation instruction, especially after the positive changes that were witnessed acoustically. However, such observations provide some interesting insights for further research. It would be interesting to explore rating scores for separate sentences, either extracted from controlled reading data (to avoid task demand effect) or from spontaneous conversational data (to avoid contextual cues for comprehensibility assessment).

Finally, the method of data collection and sound quality of spontaneous data may also account for the conflicting results between the acoustic analysis and perceived comprehensibility and accentedness. The absence of a controlled laboratory environment and supervision during data collection introduces challenges and uncontrollable environmental factors that could diminish the quality of participants' monologue production. Consequently, this may influence the judges' subjective perception of speech comprehensibility and accentedness. In contrast, the controlled reading task is less challenging, and acoustic analysis, being more objective and

sensitive to changes in intensity, F0, and duration of speech, may not align seamlessly with changes in perception scores.

In terms of the observation that comprehensibility scores are consistently higher than accentedness scores, raters' evaluations on comprehensibility may not solely rely on pronunciation. Given that each randomized audio file was an approximately 1-minute-long, prompt-guided monologue, raters could easily comprehend what participants were conveying based on contextual cues, thus leading to a higher comprehensibility score. Besides, variations in participants' expressiveness could inadvertently affect evaluation criteria on comprehensibility too, namely, a speech featuring more intricate or compelling content, or more accurate use of vocabulary and grammar might be instinctively rated higher. This aligns with the findings of Li et al. (2023) who reflected on why embodied prosodic training outperformed non-embodied training in accentedness yet failed in comprehensibility and fluency. They reinforced the notion that accentedness primarily hinges on pronunciation factors while comprehensibility encompasses additional factors like grammar and lexis. In the current evaluation task, despite explicit reminders for raters to focus exclusively on pronunciation during the assessment, one judge still raised such a concern upon completion of the evaluation task, stating:

> "I suspect I have given higher scores to the speakers with more sophisticated vocabulary and fewer pauses, even though those factors were not supposed to be evaluated. One tends to form a holistic image of the speaker's competence and then adjust the specific factors accordingly."

5.3. Discussion of the satisfaction surveys

In order to gather participants' feedback on the training course, a satisfaction survey was conducted in each group. First of all, the importance of pronunciation was acknowledged by all participants as well as the weakness of their own productions. The performed t-tests did not find any significant difference, which highlights the demand for a tailored pronunciation training course for all participants.

Regarding participants' satisfaction towards the training course, all learners from the explicit group rated the course *useful* or *very useful*, and in total 91% participants claimed they understood *much* or *all* of the rules. Seen from the implicit group feedback, most of the participants were also positive about the training and 82% rated it *useful* or *very useful*, but imitation practice was accepted by only 55% of the learners as *much helpful* or *very much helpful* for improving their pronunciation. The t-test performance reveals a significant difference between the two groups regarding the usefulness of the course ($t = 2.61$, $p = 0.02$), namely, the explicit approach was considered significantly more useful for the participants, compared with the implicit approach. This could be interpreted as meaning that an implicit approach without explicit guidance for learners might cause confusion for them and affect how much or how easily they benefit from the pronunciation instruction. The explicit approach is, by contrast, more straightforward and acceptable. Participants were instructed clearly about the rules instead of inferring them by themselves. Just as Kissling (2013) suggested, explicit instruction might be appropriate for advance learners, whereas implicit instruction is more effective for beginners who have no prior knowledge.

The results of the questionnaire go in line with other studies which suggest that EFL learners find training on pronunciation effective. In Kebboua Chaker's (2019)'s study of pronunciation instruction on word and sentence stress, 96% of the trainees reported in the questionnaire that they felt very satisfied about doing the pronunciation course, and the course was beneficial and useful for them. Likewise, students from Quesada Vázquez's (2019) study, who received explicit rhythm instruction for improving comprehensibility and fluency, also reported an overall positive impression of the pronunciation module and considered that the sessions provided both production and perception benefits. Therefore, it can be confidently asserted that language learners readily embrace pronunciation training when effective techniques are employed.

However, in ESL or EFL contexts, lack of training and insecurity on the part of the teachers often results in pronunciation being excluded from teaching syllabi completely. On the other hand, learners also feel challenged to acquire anything close

113

to nativelike pronunciation and believe that the required effort outweighs its potential benefits (Romero & Quesada Vázquez, 2023). Yet, putting aside the goal of nativelikeness, achieving a degree of comfortable intelligibility and comprehensibility with accurate prosody is realistically feasible, especially by equipping learners with adequate knowledge of segmental and suprasegmental features of the language, raising learners' awareness and providing them with intensive practice and corrective feedback.

5.4. Conclusions of the current study

Research has underscored the pivotal role of pronunciation in improving L2 learners' oral skills. A mere mastery of grammatical rules and rich vocabulary expression, in the absence of appropriate pronunciation, do not ensure an effective communication (Pardede, 2018, p. 143). Yet, the teaching of pronunciation in ESL or EFL context has been relegated to the sidelines (Derwing & Munro, 2005), lagging far behind the other aspects of second language acquisition, such as grammar, vocabulary or reading. Many language teachers find the task of teaching pronunciation daunting and express resistance to incorporating it into their pedagogical practices (Foote et al., 2011). Particularly, the instruction of suprasegmental elements (e.g. pitch accent, tone, stress, intonation) is a major part that is overlooked in classrooms.

Examining the status quo of English learning in China, despite years of dedicated efforts to English education, with compulsory lessons starting from Grade 3 of primary school, numerous learners struggle to engage in effective English conversation or make themselves understood by native speakers or English learners from diverse language backgrounds. One significant contributing factor to their problematic pronunciation is the misplacement of stress. As reported by Bu and Zhou (2021), Chinese learners of English have difficulties in distinguishing stressed, unstressed, and weakly stressed syllables. They tend to assign stress incorrectly, both within words (especially derivational and compound words) and within sentences where the nuclear stressed syllable lacks prominence. Additionally, Chinese English speakers often emphasize

unstressed syllables nearly as much as stressed syllables, lacking awareness of differences in intensity, pitch, and duration.

Therefore, this study intends to improve Chinese EFL learneers' pronunciation as far as word stress and sentence stress production are concerned. Two instructional approaches—explicit and implicit—were employed in separate groups to investigagte which approach proves more effective. The results of the acoustic analyses, comparing the pre- and the post- productions across the two groups, reveal that, while both approaches improved participants' pronunciation in stress production, when taking the instances of statistical significance into account, the improvement was more pronounced in the explicit group.

Concerning the spontaneous production data, although both groups exihibited improvements in accentedness, only the explicit group showed any improvement in comprehensibility. Nevertheless, neither group showed significant improvements in either aspect. These findings seem to provide further evidence supporting the effectiveness of implementing explicit instruction in teaching prosody within EFL contexts, especially in short-term training.

Yet, the merits of implicit instruction should not be overlooked either. Particularly, the qualitative satisfaction surveys revealed the participants' positive attitudes toward either instructional method, with an overall recognition of the usefulness of each approach. Therefore, it is anticipated that a third mode of instruction that combines both implicit and explicit modalities would be even more informative. On the one hand, teachers should be encouraged to maximize the learners' explicit knowledge of L2 features, emphasizing the nauances of similarities and differences of phonological systems between the native language and the target language (Naeini & Adni, 2017). On the other hand, immersing learners in a varied native speech corpus and encouraging them to mimic different speech styles and accents facilitates an intuitive assimilation of prosodic features (Yang, 2016). It is suggested that language teachers should overcome the intimidation of teaching pronunciation in classrooms and actively employ

diverse techniques, either explicit or implicit, to assist ESL or EFL learners in improving their pronunciation.

5.5. Limitations of the current study and suggestions for further research

Notwithstanding the results, it is worth bearing in mind certain limitations in this study. Firstly, the duration of the treatment and the sample size were less than ideal. Especially due to lack of regulations for mandatory attendance, the oral productions from participants who missed a portion of the sessions had to be discarded. Dedicating more time to implicit instruction and providing sufficient authentic input for intensive and extensive rehearsal may yield more satisfactory outcomes in the implicit group.

Secondly, the pre-test and post-test recordings were conducted remotely without supervision or ideal equipement and environment. Thus, it was difficult to avoid the impacts of external factors and distractors to maintain the quality of audio recordings consistently high, although noises were mitigated to some extent with audio editing softwares. A data collection carried out in a laboratory context would provide further validity to the results.

Thirdly, the study only examined the immediate effects of implicit and explicit teaching methods following a short-duration pronunciation training course. The long-term effects of the two methods on Chinese EFL learners' pronunciation are still left to speculation. A delayed test could demonstrate whether the observed improvements sustain over time. Consequently, variations in the effectiveness of the two methods may emerge in the extended duration. Specifically, the implicit group might catch up with the explicit group given that implicit knowledge often takes more time to develop (Ellis, 2015).

Last but not least, incorporating a third mode of instruction that combines both implicit and explicit modalities would provide further insights. This would enable us to test whether the integration of these two methods results in more efficient acquisition of L2 prosody.

References

Abdolmanafi-Rokni, S. J. (2009). A comparative study of the effect of explicit-inductive and explicit-deductive grammar instruction in EFL contexts. *Language in India*, *9*(11), 151-175.

Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh University Press.

Altmann, H. (2006). *The perception and production of second language stress: a cross-linguistic experimental study*. [Doctoral dissertation, University of Delaware]. University of Stuttgart Repository. https://www.ling.uni-stuttgart.de/institut/ifla/PDF_Upload/Heidi_Altmann/altmann-dissertation.pdf

Archibald, J. (1997). The acquisition of English stress by speakers of nonaccentual languages: Lexical storage versus computation of stress. *Linguistics*, *35*, 167-181. http://dx.doi.org/10.1515/ling.1997.35.1.167

Bailey, A. A., & Brandl, A. (2013). Incorporating pronunciation in the first-year Spanish classroom: An early intervention. In J. Levis, & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 207–223). Iowa State University.

Bardovi-Harlig, K. (1986). *Pragmatic determinants of English sentence stress*. Reproduced by the Indiana University Linguistics Club.

Beckman, M. E. (1986). *Stress and non-stress accent*. De Gruyter Mouton. https://doi.org/10.1515/9783110874020

Best, C. T., & Tyler. M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege* (pp. 13–34). John Benjamins. https://doi.org/10.1075/lllt.17.07bes

Boersma, P., & Weenink, D (2022). *Praat: doing phonetics by computer [Computer program]*. Version 6.2.06, retrieved 23 January 2022 from https://www.praat.org.

Bolinger, D. L. (1972). Accent is predictable (If you're a mind-reader). *Language*, *48*(3), 633. https://doi.org/10.2307/412039

Bovee, N., & Stewart, J. (2009). The utility of shadowing. In A. M. Stoke (Ed.), *JALT 2008 conference proceedings*. JALT.

Bian, F. (2013). The influence of Chinese stress on English pronunciation teaching and learning. *English Language Teaching*, *6*(11), 199-211. https://doi.org/10.5539/elt.v6n11p199

Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge University Press.

Brown, H. D. (2007). *Principles of language learning and teaching*. Pearson Longman.

Bu, Y., & Zhou, Z. (2021). A study of Chinese EFL learners' problems with stress acquisition. *System*, *96*, 102392. https://doi.org/10.1016/j.system.2020.102392

Busà, M. G. (2008), Teaching prosody to Italian learners of English: working towards a new approach. In C. Taylor (Ed.), *Ecolingua: The role of e-corpora in translation and language learning*. (pp. 113-126). Edizioni Università di Trieste.

Busà, M. G. (2012). The role of prosody in pronunciation teaching. A growing appreciation?. In M. G. Busà, & A. Stella (Eds.), *Methodological perspectives on second language prosody: papers from ML2P 2012* (pp. 101-106). Cleup.

Celce-Murcia, M. (1983). Teaching pronunciation communicatively. *Mextesol Journal*, *7*(1), 10-25. https://eric.ed.gov/?id=ED234602

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide*. (2nd ed.). Cambridge University Press.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. University of California Press.

Chao, Y. R. (1980). Chinese tone and English stress. In L. R. Waugh, & C. H. Van Schooneveld (Eds.), *The melody of language* (pp. 41-44). University Park Press.

Chen, J. (1984). *Spoken Chinese*. Beijing Press.

Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic cues to perception of word stress by English, Mandarin, and Russian speakers. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1468-1479.

Chun, D. (1982). *A contrastive study of the suprasegmental pitch in modern German, American English and Mandarin Chinese* [Unpublished doctoral dissertation]. University of California, Berkeley.

Chung, W. L. (2008). *The effectiveness of explicit, implicit, and noticing instruction: Mandarin speakers' perceptions and production of English sentence stress* [Unpublished doctoral dissertation]. Alliant International University.

Clark, J., & Yallop, C. (1995). *An introduction to phonetics and phonology*. Second edition. Blackwell.

Collins, B., & Mees, I. M. (2013). *Practical phonetics and phonology: A resource book for students*. Routledge.

Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect*, *18*(3), 53-70.

Crystal, D. (1969). *Prosodic systems and intonation in English*. Cambridge University Press

Cutler, A., & Clifton, C. (1984). The use of prosodic information in word recognition. In H. Bouma, & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language process* (pp. 183-96). Lawrence Erlbaum.

Dahmen, S., Grice, M., & Roessig, S. (2023). Prosodic and segmental aspects of pronunciation training and their effects on L2. *Languages*, *8*(1), 74. https://doi.org/10.3390/languages8010074

DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty, & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Blackwell. http://dx.doi.org/10.1002/9780470756492.ch11

Derwing, T. M. (2018). The efficacy of pronunciation instruction. In O. Kang, R. I. Thomson, & J. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 320-334). Routledge.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, *39*(3), 379–397. https://doi.org/10.2307/3588486

Derwing, T. M., Munro, M. J., & Wiebe, G. (1997). Pronunciation instruction for "fossilized" learners: Can it help?. *Applied Language Learning*, *8*, 217-236.

Derwing, T. M., Munro, M. J., & Wiebe, G. E. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *39*, 393-410.

Derwing, T. M., & Rossiter, M. J. (2002). ESL learners' perceptions of their pronunciation needs and strategies. *System*, *30*(2), 155-166.

Derwing, T. M. & Munro, M. J., (2014). Once you have been speaking a second language for years, it's too late to change your pronunciation. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 34–55). University of Michigan Press.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). John Benjamins.

Duanmu, S. (2000). *The phonology of Standard Chinese*. Oxford University Press.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America. 80*, 402–415.

Ellis, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Multilingual Matters.

Ellis, R. (2015). *Understanding second language acquisition, 2nd edition*. Oxford university press. DOI:10.1016/j.system.2017.07.013

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

Flege, J. E., & MacKay, I. R. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, *26*, 1–34.

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada journal*, 1-22.

Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, *3*(1), 34-56.

Frost, D., & Picavet, F. (2014). Putting prosody first-Some practical solutions to a perennial problem: The Innovalangues project. *Research in Language*, *12*(3), 233-243.

Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765–768.

Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, *1*(2), 126-152.

Fry, D. B. (1965). The dependence of stress judgments on vowel formant structure. In E. Zwirner, & W. Bethge (Eds.), *Proceedings of the 5th international congress of phonetic sciences*, 306-311. Karger Publishers.

Fu, Q. J., & Zeng, F. G. (2000). Identification of temporal envelope cues in Chinese tone recognition. *Asia Pacific Journal of Speech, Language and Hearing*, *5*(1), 45-57.

Fullana, N. (2006). The development of English (FL) perception and production skills: Starting age and exposure effects. *Age and the rate of foreign language learning*, *19*, 41-64.

Gardner, R. C., & Lambert, W. E. (1972). *Attitudes and motivation in second-language learning*. Newbury house.

Ghorbani, M. R., Neissari, M., & Kargozari, H. R. (2016). The effect of explicit pronunciation instruction on undergraduate EFL learners' vowel perception. *Language and Literacy*, *18*(1), 57-70.

Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis, & K. LeVelle (Eds.). *Proceedings of the 4th pronunciation in second language learning and teaching conference. Aug. 2012.* (pp. 194–206). Iowa State University.

Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard*, *3*(1), 20170007.
DOI: http://doi.org/10.1515/lingvan-2017-0007

Guo, X. (2022). Acoustic correlates of English lexical stress produced by Chinese dialect speakers compared to native English speakers. *Frontiers in Psychology*, *13*, 796252. https://doi.org/10.3389/fpsyg.2022.796252

Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*(2), 201–223.

Hamada, Y. (2018). Shadowing for pronunciation development: Haptic-shadowing and IPA-shadowing. *Journal of Asia TEFL*, *15*(1), 167.

Hamzah, M. H. (2014). The role of explicit phonetic instruction in pronunciation teaching in ESL settings [Conference presentation]. *Prosiding seminar Antarabangsa Kelestarian Insan 2014 (INSAN 2014)*, Batu Pahat, Johor.

Hara, Y., Kawahara, S., & Feng, Y. (2014). The prosody of enhanced bias in Mandarin and Japanese negative questions. *Lingua*, *150*, 92-116.

He, S. F. (2002). *Contrastive analysis of English and Chinese*. Shanghai Foreign Language Education Press.

Hewings, M. (2007). *English pronunciation in use advanced book with answers, 5 audio CDs and CD-ROM*. Cambridge University Press.

Hismanoglu, M. (2009). The pronunciation of the inter-dental sounds of English: an articulation problem for Turkish learners of English and solutions. *Procedia-Social and Behavioral Sciences*, *1*(1), 1697-1703.

Housen, A. & Pierrard, M. (2005). Investigating instructed second language acquisition. In A. Housen, & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 1–30). De Gruyter Mouton.

Howatt, A. P. R., & Widdowson, H. G. (2004). *A History of English Language Teaching* (2nd ed.). Oxford University Press.

Hsieh, K. T., Dong, D. H., & Wang, L. Y. (2013). A preliminary study of applying shadowing technique to English intonation instruction. *Taiwan Journal of Linguistics*, *11*(29), 43-66.

Hu, F. (2017). A study on Chinese EFL learners' phonetic obstacles to listening comprehension. *Journal of Language Teaching and Research*, *8*(2), 404-410. DOI: http://dx.doi.org/10.17507/jltr.0802.25

Hyman, L. M. (1977). On the nature of linguistic stress. In L. M. Hyman (Ed.), *Studies in stress and accent. SCOPIL* 4. (pp. 37-82). Department of Linguistics, University of Southern California.

Jenkins, J. (2000). *The phonology of English as an international language: New models, new Norms, new Goals*. Oxford University Press.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics*, *23*(1), 83–103. https://doi.org/10.1093/applin/23.1.83

Juffs, A. (1990). Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *International Review of Applied Linguistics*, *28*(2), 99-118. http://dx.doi.org/10.1515/iral.1990.28.2.99

Jung, Y. J., & Rhee, S. C. (2018). Acoustic analysis of English lexical stress produced by Korean, Japanese and Taiwanese-Chinese speakers. *Phonetics and Speech Sciences*, *10*(1), 15-22.

Kadota, S. (2019). *Shadowing as a practice in second language acquisition: Connecting inputs and outputs*. Routledge.

Kebboua Chaker, N. (2019). *Integrating information technology in the teaching/learning of english pronunciation in the classroom: designing and implementing an online course to teach word and sentence stress to tertiary level students* [Doctoral dissertation, Universitat Rovira i Virgili]. URV Repository.

Kelly, L. G. (1969). *25 Centuries of language teaching: An inquiry into the Science,*

Art, and development of language teaching methodology, 500 BC-1969. Newbury House Publishers.

Kenworthy, J. (1987). *Teaching English pronunciation*. Longman.

Keyworth, P. R. (2015). An acoustic phonetic study on the production accuracy of lexical stress contrasts in Arabic and mandarin-accented English. *Journal of the Acoustical Society of America*, *137*, 2271–2272. doi: 10.1121/1.4920296

Khanbeiki, R., & Abdolmanafi-Rokni, S. J. (2015). The effect of explicit vs. implicit instruction on the learnability of English consonant clusters by Iranian learners of English. *Advances in Language and Literary Studies*, *6*(2), 103-112.

Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal, 97*(3), 720-744.

Koike, Y. (2014). Explicit pronunciation instruction: Teaching suprasegmentals to Japanese learners of English. In N. Sonda, & A. Krause (Eds.), *JALT 2013 conference proceedings* (361-374). JALT.

Krashen, S., & Terrell, C. (1983). *Natural Approach*. Pergamon Press.

Kucukoglu, H. (2012). Sentence stress and learning difficulties of ELT teachers: A case study. *Procedia-Social and Behavioral Sciences*, *46*, 4065-4069.

Kuiper, K., & Allen, W. S. (1996). *An introduction to English language*. Macmillan.

Lacabex, E. G., & Gallardo-del-Puerto, F. (2020). Explicit phonetic instruction vs. implicit attention to native exposure: phonological awareness of English schwa in CLIL. *International Review of Applied Linguistics in Language Teaching*, *58*(4), 419-442.

Ladd, D. R. (1996). *Intonational phonology*. Cambridge: Cambridge University Press.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge. University Press. https://doi-org.sabidi.urv.cat/10.1017/cbo9780511808814

Laver, J. (1994). *Principles of phonetics*. Cambridge University Press.

Lee, K. (2013). Sentence stress in information structure. *Oenoehag, 27*, 3-30.

Levis, J. M. (2015). Learners' views of social issues in pronunciation learning. *Journal of Academic Language and Learning*, *9*, A42-A55.

Levis, J., & Sonsaat, S. (2017). Pronunciation teaching in the early CLT era. In O. Kang, R. Thomson, & J. Murphy (Eds.), *The Routledge handbook of English pronunciation* (pp. 267-283). Routledge.

Li, K. (2008). Contrastive focus structure in Mandarin Chinese. In M. Chan & H. Kang (Eds.), *Proceedings of the 20th North American conference on Chinese linguistics* (*NACCL-20*) (pp. 759-774). The Ohio State University.

Li, P., Baills, F., Baqué, L., & Prieto, P. (2023). The effectiveness of embodied prosodic training in L2 accentedness and vowel accuracy. *Second Language Research*, *39*(4), 1077-1105. https://doi.org/10.1177/02676583221124075

Liberman, M. Y. (1975). *The intonational system of English* [Doctoral dissertation, Massachusetts Institute of Technolog]. MIT Repository.

Lin, M. C. (1988). Putonghua shengdiao de shengxue xingzhi yu zhijue zhengzhao [acoustic properties and perceptual cues of mandarin tones]. *Zhongguo Yuwen*, *204*, 437–445.

Lord, G. (2005). (How) Can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, *88*(3), 557–567. https://doi.org/10.2307/20063159

Macdonald, D., Yule, G., & Powers, M. (1994). Attempts to improve English L2 pronunciation: The variable effects of different types of instruction. *Language Learning*, *44*(1), 75-100.

Mackay, I. (1985). *Introducing Practical Phonetics*. Crane Inc.

Minhong, Y., & Ailun, L. (2006). The effect of implicit learning on the learning of a second language pronunciation. *Psychological Science*, *3*, 26.

Mott, B. L. (1991). *A course in phonetics and phonology for Spanish learners of English*. Promociones y Publicaciones Universitarias.

Morley, J. (1987). *Current perspectives on pronunciation*. TESOL.

Munro, M., & Derwing, T. (1995). Processing time, accent, and comprehensibility in the perception of native and foreign-accented speech. *Language and Speech*, *38*, 289–306.

Murphy, J. M., & Baker, A. A. (2015). History of ESL pronunciation teaching. In M. Reed, & J. M. Levis (Eds.), *The Handbook of English Pronunciation* (pp. 36-65). Wiley-Blackwell.

Naeini, B., & Adni, Z. (2017). The impact of explicit vs. implicit instruction on pronunciation intelligibility among Iranian EFL learners. *Journal of Studies in Learning and Teaching English*, *6*(2), 103-123.

Nguyen, A.T. T., Ingram, J., & Pensalfini, R. (2008). Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics*, *36*, 158–190.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language learning*, *50*(3), 417-528.

Odlin, T. (1989). *Language transfer* (Vol. 27). Cambridge University Press.

Papachristou, V. (2011). Explicit vs. implicit pronunciation teaching to Greek children: The case of the acquisition of English vowels. *Selected papers from the 19th ISTAL*, 371-381. https://doi.org/10.26262/istal.v19i0.5515

Pardede, P. (2018). Improving EFL students' English pronunciation by using the explicit teaching approach. *Journal of English Teaching*, *4*(3), 143–155.

Peltekov, P. (2020). The effectiveness of implicit and explicit instruction on German L2 learners' pronunciation. *Die Unterrichtspraxis/teaching German*, *53*(1), 1-22.

Pennington, M. C., & Rogerson-Revell, P. (2019). English pronunciation teaching and research. *Research and Practice in Applied Linguistics*: *Contemporary Perspectives*. Palgrave Macmillan.

Pica, T. (1996). The essential role of negotiation in the communicative classroom. *JALT Journal*, *18*(2), 241-268.

Prator, C. H., & Robinett, B. W. (1985). *Manual of American English pronunciation* (4th ed.). Holt, Rinehart and Winston.

Quesada Vázquez, L. (2019). *The introduction of rhythm instruction in the English as a foreign language classroom to improve the comprehensibility and fluency of*

English for specific purposes students [Doctoral dissertation, Universitat Rovira i Virgili]. URV Repository.

Ringbom, H. (2007). *Cross-linguistic similarity in foreign language learning* (Vol. 21). Multilingual Matters.

Roach, P. (2009). *English phonetics and phonology: A practical course*. Cambridge University Press.

Roccamo, A. (2015). Teaching pronunciation in just ten minutes a day: A method for pronunciation instruction in first-semester German language classrooms. *Die Unterrichtspraxis/Teaching German*, *48*(1), 59–83. https://doi.org/10.1111/tger.10181

Romero, J., & Quesada Vázquez, L. (2023). Objective versus perceived proficiency in pronunciation assessment of phonetic detail in advanced L2 English learners. In R. Skarnitzl, & J. Volín (Eds.), *Proceedings of the 20th International Congress of Phonetic Sciences* (pp. 4239–4243). Guarant International.

Ruellot, V. (2011). English in French print advertising from 1999 to 2007. *World Englishes*, *30*(1), 5-20.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: an instructed SLA approach to L2 phonology. *Language Awareness*, *20*(1), 45–59. https://doi.org/10.1080/09658416.2010.540326

Saito, Y., & Saito, K. (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research*, *21*(5), 589-608.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge University Press.

Shamiri, H., & Farvardin, M. T. (2016). The effect of implicit versus explicit corrective feedback on intermediate EFL learners' speaking self-efficacy beliefs. *Theory and*

Practice in Language Studies, *6*(5), 1066–1075. https://doi.org/10.17507/tpls.0605.22

Sharma, L. R. (2021). Significance of teaching the pronunciation of segmental and suprasegmental features of English. *Interdisciplinary Research in Education*, *6*(2), 63-78. https://doi.org/10.3126/ire.v6i2.43539

Shen, X. S. (1993). Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech*, *36*(4), 415-433.

Spaai, G. W., & Hermes, D. J. (1993). A visual display for the teaching of intonation. *Calico Journal*, *10*(3), 19-30.

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

Terken, J., & Hermes, D. (2000). The perception of prosodic prominence. In M. Horne (Ed.), *Prosody: Theory and experiment. Studies presented to Gösta Bruce* (pp. 89-127). Kluwer.

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

Trask, R. L. (1996). *A dictionary of phonetics and phonology*. Routledge.

Třísková, H. (2016). De-stressed words in Mandarin: Drawing parallel with English. *Integrating Chinese linguistics research and language teaching and learning*, 121-144.

Trofimovich, P., & Baker, W. (2006). Learning second language suprasegmentals: Effect of L2 experience on prosody and fluency characteristics of L2 speech. *Studies in second language acquisition*, *28*(1), 1-30.

Tsiartsioni, E. (2011). Can pronunciation be taught? Teaching English speech rhythm to Greek students. *Selected papers on theoretical and applied linguistics*, *19*, 447-458.

Tuan, L. T., & An, P. T. V. (2010). Teaching English rhythm by using songs. *Studies in Literature and Language*, *1*(2), 13-29.

Vigliano, D., Yoshimoto, K., & Pellegrino, E. (2016). A self-imitation technique for the improvement of prosody in L2 Italian. *Proceedings of 22nd Annual Meeting of the Association for Natural Language Processing* (pp. 1189-1192). Tohoku, Japan.

Wang, Q. (2008). *Perception of English stress by Mandarin Chinese learners of English: An acoustic study* [Doctoral dissertation, University of Victoria]. University of Victoria Repository. http://hdl.handle.net/1828/1282

Wang, M., Koda, K., & Perfetti, C. A. (2003). Alphabetic and nonalphabetic L1 effects in English word identification: A comparison of Korean and Chinese English L2 learners. *Cognition*, *87*(2), 129-149.

Wang, Y., & Chu, M. (2006). An experimental study on the assignment of focus accent in Mandarin. *Proc. Speech Prosody 2006*, paper 180. https://www.isca-speech.org/archive/speechprosody_2006/wang06f_speechprosody.html

Xu, Y. M. (2019). *Contrastive studies of sound systems in English and Chinese*. Foreign language Teaching and Research Press.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of phonetics*, *27*(1), 55-105.

Yakut, İ. (2020). Promoting the correct production of English sounds in extensive reading-circle classes: Explicit vs. implicit pronunciation training. *Eurasian Journal of Applied Linguistics*, *6*(1), 101-118.

Yang, L. (2016). Optimizing pronunciation and prosody teaching in second language learning. In *ISAPh 2016 International Symposium on Applied Phonetics* (pp. 68–73). Nagoya, Japan. DOI:10.21437/ISAPh.2016

Zhang, Y, Nissen, S. L, & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America, 123*, 4498–4513.

Zielinski, B.W. (2008). The listener: no longer the silent partner in reduced intelligibility. *System 36*(1), 69–84.

Zielinski, B. (2015). The segmental/suprasegmental debate. In M. Reed, & J. M. Levis

(Eds.), *The handbook of English pronunciation* (pp. 397-412). Wiley Blackwell.

https://doi-org.sabidi.urv.cat/10.1002/9781118346952.ch22

Appendix A. Pretest/Posttest

Please follow the instructions and record your reading in a quiet environment. For some parts, please record the words/sentences in one audio, as indicated in (brackets). For some, you may need to divide the sentences into several recordings.

Task 1 Reading

● Word reading

1) curiosity   2) maintain   3) analysis   4) satisfactory

5) agriculture   6) analyze   7) cooperate   8) immigrant

● Discourse reading

Part 1

A: Tom lost his key.

B: I saw it on the table.

A: Have you taken it?

B: Yes, I put it in my bag.

A: Can you give it to him?

B: Sure. Where is he?

A: He is playing on the playground.

B: Ok, I'll meet him after class.

Part 2

A: I can't find my dress.

B: What kind of dress?

A: It is a beach dress.

B: What color beach dress?

A: It is white. White with stripes.

B: There was a white dress with stripes in the car.

A: Which car?

B: The one I sold.

Part 3

1. He is a doctor and I am a nurse.

2. I love playing badminton rather than watching it.

3. Is he traveling here?   No, he is living here.

4. I prefer a car to a bike.

5. He cleaned the room, not me.

6. It is cold in the morning and hot in the afternoon.

Part 4

1. A: How was the movie?

   B: It was very boring.

2. A: Why is he not here?

   B: He is always late.

3. A: How is your holiday?

   B: It is absolutely interesting.

4. A: How do you like the view?

   B: It is really pretty.

5. A: Do you like the ice cream?

   B: Yes, it is so tasty.

6: Can you go to the meeting?

   I told you I can't go. I feel terribly sorry.

Task 2 free-talk (for pretest)

Look at the pictures and answer the following questions about different living places.

➡   Do you like living in the city or the countryside?

➡   What are the benefits and drawbacks of living in the city and countryside?

You can talk up to 1 minute. Please record yourself without any noise.

PS: To know your real speaking level, please don't write down your answer and read

it.

Due to copyright issue, the city picture prompt is not presented.

Task 2 free-talk (for posttest)

The following picture presents two common methods of shopping.

➥ What are the pros and cons of each shopping habit?

➥ Which do you prefer?

You can talk up to 1 minute. Please record your response without any noise.

PS: To know your real speaking level, please don't write down your answer and read it.

Appendix B. Explicit training course materials

Captures from the training PowerPoint

TABLE OF CONTENT

Lesson 1 Syllables and word stress
Lesson 2 Sentence stress
Lesson 3 A thought group & tonic stress
Lesson 4 Stress on new information
Lesson 5 Contrastive tress
Lesson 6 Emphatic stress
Lesson 7 Non-prominent words
Lesson 8 Review

Module 1: Syllables and word stress

OVERVIEW

Session 1

- Chinese and English are very different in terms of pronunciation rules.

Chinese With tones; No word stress **VS** **English** With word stress; No tones

- What is tone?
  There are 5 tones on one Chinese syllable, but English syllables have no tones.
  e.g. ma 5 tones （corresponding Mandarin Chinese:妈 嘛 马 骂 吗）
- English has word stress, but Chinese doesn't.
  e.g. mother /ˈmʌðə/ computer /kəmˈpjuːtə/

SYLLABLES

Every syllable of every word must have at least one vowel sound. A vowel can stand alone in a syllable, as in **u•nit**.
It can also be surrounded by consonants, as in jet, nap•kin, and fan•tas•tic.

PRACTICE

Please count how many syllables each word has.

1. small   2. electric   3. unsatisfactory   4. gift

5. cooker   6. kindergarten   7. analysis   8. analyze

QUESTION

How many syllables does a Chinese word have?

One Chinese word only has one syllable, e.g. 妈 （ma）, 包 （bao）.

ENGLISH WORD STRESS

- Can you hear any difference from the two recordings?

- Routine

WORD STRESS

- Each English word has one or more syllables but only one **stress**. The stress can be shown in the word's phonetic symbol by a semicolon on the upper left of the syllable ˈ (**main stress**) . Some long words also have a secondary stress (with symbol ˌ on the lower left of the syllable), softer than the main stress.
- e.g. **par**ty /ˈpɑːti/

    kindergarten /ˈkɪndəˌgɑːt(ə)n/

    electricity /ˌɪlɛkˈtrɪsɪti /

Another common way to show word stress:

O o   O o Oo   o O
party   kindergarten   dessert

WORD STRESS

A stressed syllable

1. is louder, (e.g. com PU ter)
2. is longer, (com p-u-ter)
3. has a higher pitch (It has a change in pitch from the syllables coming before and afterwards)
4. contains a full voiced vowel, said more clearly.   (Roach, 1983: 73)

computer

Pitch

Intensity

Duration

com   pu   ter
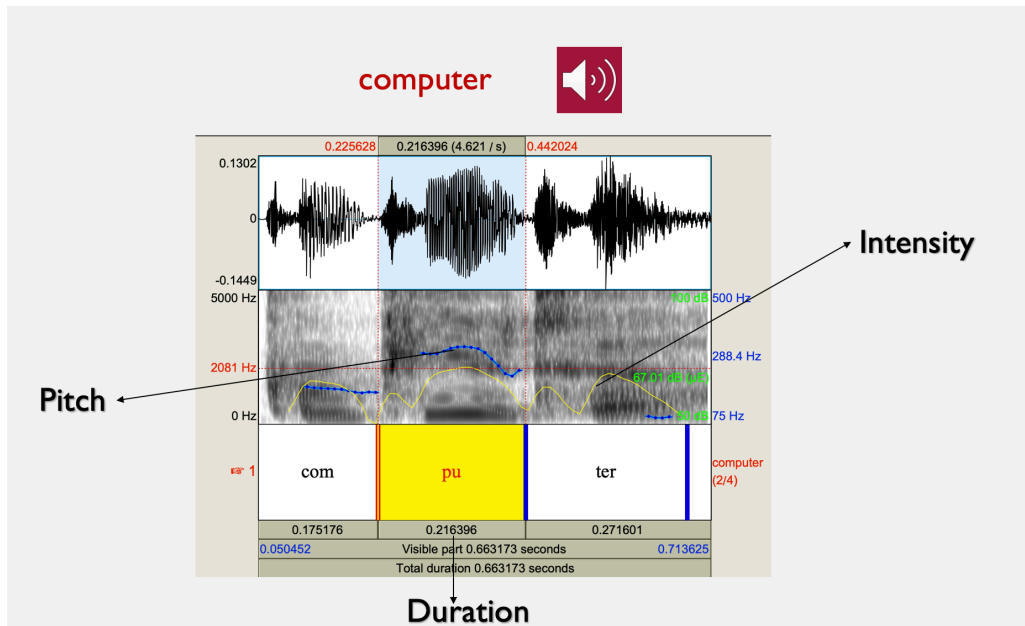
PRACTICE

- Present
- Experiment
- Spectacular
- Documentary
- Supplement

MATERIALS
TO BE RECORDED

Read the words and record yourself.

Satisfactory

Analysis

Inconclusive

Harmonious

Chinese

Confidence

Module 2: Sentence stress

SENTENCE STRESS

Session 2

- Listen to the recordings, can you hear any difference? What is the difference?

The children will be swimming in the river.

SENTENCE STRESS

Sentence stress is what gives English its **rhythm** or "beat". Some words in a sentence are stressed, some are not.

Most sentences have 2 basic types of words: content words *vs* function words

Content words: the important words which carry the meaning;

Function words: small and simple words which make a sentence correct grammatically.

- e.g.   A **bag** of **rice**
- A **nice family**
- I **can't**!
- She **has** a **book.**

SENTENCE STRESS

If you remove the function words from a sentence, you will probably still understand the sentence. But if you remove content words, you lose the meaning.

- e.g.   **sell**   **car**   **gone**   **France**

If you add a few more words, you can make the sentence complete and correct.

Will you **sell** my **car** because I've **gone** to **France**

English rhythm: the content words are stressed but function words are not.

CONTENT WORDS *VS* FUNCTION WORDS

Content words: carry semantic meaning	Function words: convey grammatical functions
<u>Nouns</u>: boy, house, etc.	<u>Auxiliary verbs</u>: does, have, etc.
<u>Adjectives</u>: big, strong, etc.	<u>Conjunctions</u>: and, but, than, etc.
<u>Verbs</u>: run, flow, etc.	<u>Prepositions</u>: on, in, etc.
<u>Adverbs</u>: slowly, carefully, etc.	<u>Articles and determiners</u>: a/an, the, some, etc.
<u>Negation words</u>: not, never, etc.	<u>(indefinite) pronouns</u>: she, he, some-/anyone, etc.
<u>Interrogative words</u>: when, how, etc.	<u>Possessive pronouns</u>: his, my, etc.
<u>Demonstratives</u>: this, that, etc.	<u>each other, one another</u>
<u>Numbers</u>: one, two, third, etc.	*There* used to start a sentence

PRACTICE

- Underline the stressed words in each sentence.

- She saw me yesterday.
- Could you check my homework?
- I made my coffee at home
- It's easy with good habits.
- His book was lost in the mall.

SENTENCE STRESS

Chinese Syllable-timed	**vs**	English Stress-timed

- **Syllable-timed**: the time to pronounce each syllable is roughly the same.
  e.g. wo-yao-qu-shang-xue   I need to go to school.
- **Stress-timed**: the time between the stressed syllables is roughly the same.
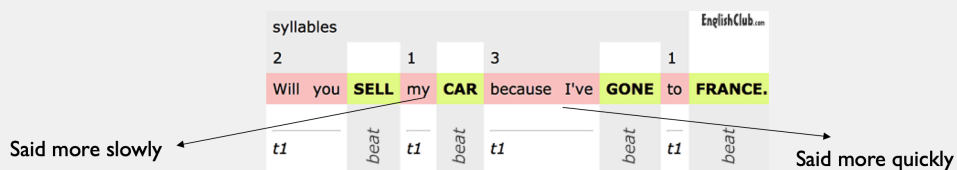
PRACTICE

- Dogs chase cats.
- The dogs chase cats.
- The dogs chase the cats.
- The dogs will chase the cats.
- The dogs will be chasing the cats

PRACTICE

Adds beats to your reading.

- She saw me yesterday.
- Could you check my homework?
- I made my coffee at home
- His book was lost in the mall.
- She bought the coffee at the bar.

MATERIALS TO BE RECORDED

Please read the sentences and record yourself clearly.

1. I'll come tomorrow.
2. She carries the wrong bag.
3. The children will be swimming in the river.
4. I want to buy a car.
5. I hope to visit my grandparents.
6. I went out for shopping.

Module 3: A thought group and tonic stress

Session 3

A THOUGHT GROUP

- Can you add more information to make the sentence longer?

- He is waiting…

A THOUGHT GROUP

- A thought group (or an intonation unit) is a group of several words that form a unit of meaning.
- A sentence can be divided into one or several thought groups.
- We can have a short pause at the end of each phrase.

e.g. He is waiting.

He is waiting / for his friend.

He is waiting / for his friend / at the station.

LISTEN AND BREAK THE UNITS

1. She'd left when she had a baby and then decided not to go back although the job had been kept open for her

2. Tom dear where's the advert for this calculator because I don't know the address and I don't know who I've got to make the cheque payable to

PRACTICE

Can you divide the sentences into thought groups?

e.g. She is a teacher.

- A beautiful woman and her dog entered the room.
- When I woke up I didn't even realise what time it was.
- Because he was ill I didn't expect him to come to work.
- When I was in the canteen I tripped over a plate on the floor but my friend caught me in her arms.

KEY

- She is a teacher. ⟸ Don't break up a short sentence.
- A beautiful woman / and her dog / entered the room.
- When I woke up / I didn't even realise / what time it was.
- Because he was ill / I didn't expect him to come to work.
- When I was in the canteen , I tripped over a plate on the floor / but my friend caught me. in her arms.

If you want to speak faster, you can break the sentence into longer groups.
If you want to speak slower, you can break the sentence into more shorter groups.

TONIC STRESS

- Each thought group or sentence has a tonic stress, also called 'nucleus', which receives the most prominence (other content words may also be stressed too but not that strong) . Tonic stress carries the most important information and normally falls on the **last content word**.
- E.g. He's **WAI**ting
- He's **wai**ting for his **FRIEND**
- He's **wai**ting for his **friend** at the **STA**tion

- p.s. remember only the **main stressed syllable** of the final content word carries the tonic stress, not the whole word.

142

PRACTICE

Break each sentence when needed and find the tonic stress.

- She is a teacher.
- A beautiful woman and her dog entered the room.
- When I woke up, I didn't even realise what time it was.
- because he was ill, I didn't expect him to come to work.
- When I was in the canteen, I tripped over a plate on the floor, but my friend caught me in her arms.

KEY

- She is a **TEA**cher.
- A <u>beautiful</u> woman and her **DOG** / <u>entered</u> the **ROOM**.
- <u>When</u> I **WOKE** up / I <u>didn't</u> even **REA**lise <u>what</u> **TIME** it was.
- Because he was **ILL** / I <u>didn't</u> <u>expect</u> him to <u>come</u> to **WORK**.
- When I was in the can**TEEN**/ I <u>tripped</u> over a <u>plate</u> on the **FLOOR** / but my <u>friend</u> <u>caught</u> me in her **ARMS**.

MATERIALS TO BE RECORDED

Please read the sentences and record yourself clearly.

- 1. Do it!
- 2. I am younger than him.
- 3. We often hang out on Sunday.
- 4. Dogs eat bones.
- 5. I need to go to the station.
- 6. He plays games on his computer.

Module 4: Stress on new information

Session 4

TONIC STRESS SHIFT

- The tonic stress can be shifted from the final content position to other words, even function words, in the following three situations:

  New information stress
  Contrastive stress
  Emphatic stress

NEW INFORMATION STRESS

- The tonic stress normally falls on the last content word if all information is new.
- If the last content word contains the old information in a conversation, the stress will shift to the proceeding content word or even the function word to <u>focus on the new information</u>.

- e.g. 1. A: What color is your **CAR**?
            B: It is **BLUE**, **DARK** blue.
       2. A: Where are you **GO**ing?
            B: To the **SU**permarket. Where are **YOU** going?
            A: I am **AL**so going to the supermarket.

NEW INFORMATION STRESS

In a response to a Wh-question, the information supplied is normally stressed too, because it is naturally the new information in the conversation.

- <u>What's</u> your **NAME**?

  My name is **TOM**.

- <u>Where</u> do you **COME** from?

  I'm from **CHINA**.

- <u>What's</u> your **JOB**?

  I work as a **TEA**cher.

- <u>What</u> do you **like**?

  I like to play **BAS**ketball.

144

PRACTICE

• Match the questions with the appropriate response.

1. a) Have you ever been to Spain before?   i) I worked in SPAIN.

   b) Where did you work?   ii) I WORKED in Spain.

2. a) That mobile looks familiar.   i) It's YOUR phone.

   b) What is making that noise?   ii) It's your PHONE.

3. a) Who is talking on the phone?   i) Lucy is making a CALL.

   b) What is Lucy doing?   ii) LUcy is making a call.

4. a) She works at St. Mary's, doesn't she?   i) She is a reCEPtionist at that hospital.

   b) She works as a receptionist nearby, doesn't she?   ii) She is a receptionist at the HOSpital.

PRACTICE

• A: Where is she from?

• B: She is from England. Where are you from?

• A: I am from China.

• A: How do you go to your office?

• B: I drive to the office.

• A: How long does it take?

• B: It's a 10-minute drive.

RECORDING MATERIALS

1. -Which bag do you want to buy?   -The yellow bag.

2. -Do you want to go to Beijing?   - Yes, it's my dream city.

3. -How do you go to school?   -I walk to school.

4. -Where is your friend from?   -She is from California.

5. -How old is she?   -She is 9 years old.

6. -A: I´ve lost my hat.

   B: What kind of hat?

   A: It was a sun hat.

Module 5: Contrastive stress

CONTRASTIVE STRESS

Session 5

- Contrastive stress: when we are making a contrast with something previously said or correct something said previously, we stress the corresponding words.

  e.g.

  A: Can I have some of this **CAKE** now?

  B: It is for to**MOR**row's party, not for to**day's** meal.

  'today' is contrasted with 'tomorrow' for correction

CONTRASTIVE STRESS

- Contrastive stress is used to point out the difference between one object and another. It tends to be used with determiners, such as 'this', 'that', 'these' and 'those'.

  e.g.

- I think I prefer **THIS** car.
- Do you want **THESE** or **THOSE** books?

PRACTICE I

- Can you find the prominent words/syllables for contrasts?

1. A: Your brother's a doctor, isn't he?   B: No, my father's a doctor.
2. I think it's more important to have comfortable clothes than stylish ones.
3. A: I'll see you in the office on Friday.   B: I work at home on Fridays.
4. Watching football is much better than watching basketball.
5. Watching football is more interesting than playing football.

One Sentence, Different Meanings

- Are you going to eat **THAT**?

- Are you going to **EAT** that?

- Are **YOU** going to eat that?

- **ARE** you going to eat that?

PRACTICE 2

- Read the answer for each question. Are there any difference?

    –Who bought a computer yesterday?

    -Jane bought a computer yesterday.

    -How did Jane get the computer yesterday?

    -Jane bought the computer yesterday.

    -What did Jane buy yesterday?

    -Jane bought a computer yesterday.

    -When did Jane buy a computer?

    -Jane bought a computer yesterday.

    -Who did Jane buy the computer with?

    - Jane bought the computer with her parents.

PRACTICE 3

Find the most prominent word from each sentence in blue.

1. A: Why don´t you try keeping a spider as a pet?

   B: //I´ve always been terrified of spiders.//

2. A: Is there anything that really terrifies you?

   B: //I´ve always been terrified of spiders.//

3. A: How long have you been frightened by spiders?

   B: //I´ve always been terrified of spiders.//

4. A: Who is always been frightened by spiders?

   B: // I´ve always been frightend by spiders.//

MATERIALS TO BE RECORDED

1. It's my book, not yours.

2. Do you want to stay or leave?

3. I prefer that color.

4. Do you enjoy watching tennis or playing tennis?

5. Do you want to go to school or go home?

6. -Who cooked the dinner? -Your sister cooked it.

Module 6: Emphatic stress

EMPHATIC STRESS

Session 6

- If you want to emphasize a certain word particularly, you can change the stress from the final tonic stress to other content word, such as an adjective.

  e.g.   That was a difficult **test**. - *Standard statement*

          That was a **difficult** test. - *Emphasizes how difficult the test was*

- *Or certain words can be used to show strong emotions.*

  e.g.   This dish is **so** tasty. – to show the strong emotion.

PRACTICE

- 1. My teacher went there by herself.
  2. The view is terrific.
  3. It's my own house.
  4. I spent a lot buying it.

MORE EXAMPLES

- Some adverbs and modifiers are emphatic by nature.

e.g. Indeed, utterly, absolute, terrific, tremendous, awful, always, great, really, truly, very, quite, entirely, extremely, only, own, alone, surely…

Can you create scenarios and make sentences with them, read aloud in class?

MATERIALS TO BE RECORDED

- 1. The watch is too expensive.
- 2. The food is awful.
- 3. It is extremely hot in summer.
- 4. It is my only chance.
- 5. He is surely fine.
- 6. She is always angry.

Module 7: Non-prominence on empty words

NON-PROMINENCE ON EMPTY WORDS

Session 7

- Some final content words are empty without being stressed because their meaning is obvious from what has been said before.

e.g.

I must get my **HAIR** cut?

She's got an **E**ssay to write.

First I've got some **SHO**pping to do.

> 'running', 'cut', 'to write' and 'to do' are predictable meanings in these contexts.

- Can you come up with other examples like this?

PRACTICE I

- She's made a name for herself.
- We don't have any food left.
- I need to study hard from now.
- How is your work going?

NON-PROMINENCE ON EMPTY WORDS

- Vague expressions are used at the end of a speech unit when we don't need to be precise. They are usually non-prominent, following a prominent word.

e.g. I bought some CHEESS and stuff.

My car is always breaking down. I HATE the thing.

How long will it take? About a WEEK or so.

We could meet about TEN or whenever.

> 'the stuff', 'the thing', 'or so' and 'or whenever' are all vague expressions.

PRACTICE 2

- She went out with Ken or somebody.
- Let me know if you need help or anything.
- Do you need water or something?
- I love hiking, cycling and these sorts of things.
- It lasts two hours or so.
- I'll pick you up at your place.

MATERIALS TO BE RECORDED

- 1. Let's have the meeting at my place.
- 2. How's your father doing?
- 3. He lives along the coast or somewhere.
- 4. I don't know where to buy the stuff.
- 5. I have to arrive at 3 o'clock.
- 6. I've got some water to drink.
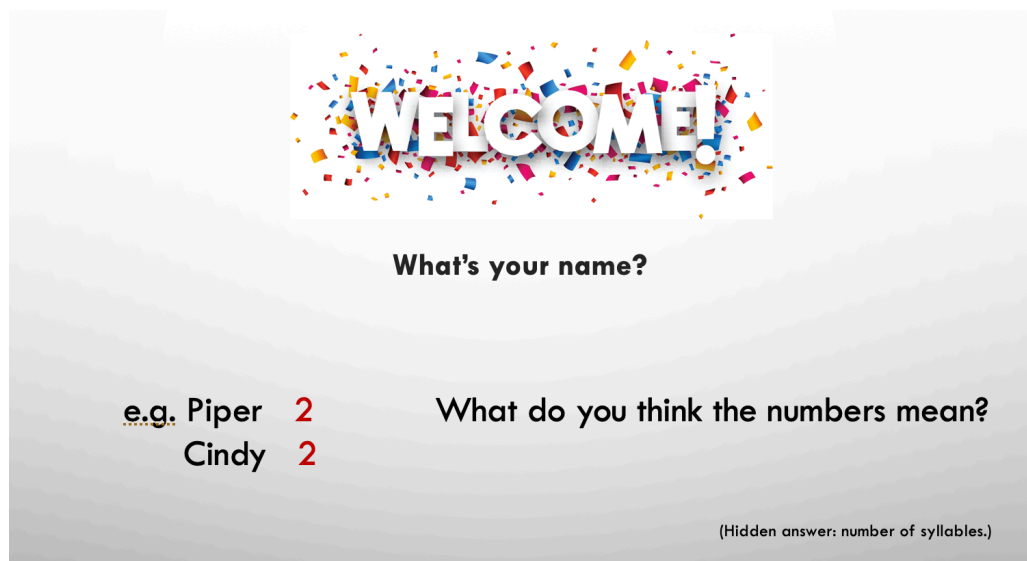
Module 8 is a review of the previously presented content.

Appendix C. Implicit training course materials

Module 1: Syllables and strength on words

LESSON 1

THE BEAT

PART 1

- LISTEN TO THE AUDIO AND SAY HOW MANY PARTS EACH WORD HAS.

- SPELL THE WORDS

Forget   experiment   fog   marathon   popularity

- REPEAT AFTER THE AUDIO

PRACTICE

- LISTEN TO THE AUDIO AND CATEGORIZE THE WORDS BASED ON THEIR NUMBER OF PARTS.
CLAP THE WORDS.

1. small   2. electric   3. unsatisfactory   4. gift

5. cooker   6. kindergarten   7. analysis   8. analyze

1   2   3   4   5

WHAT DO WE CALL THESE PARTS?

SYLLABLE

e.g. cooker
 analysis
 expert

Q1： How many syllables are they?
Q2： What attribute to a syllable?

Every syllable of every word must have at least one vowel sound. A vowel can stand alone in a syllable, as in **u•nit**.
It can also be surrounded by consonants, as in jet, nap•kin, and fan•tas•tic.

PRACTICE

• WITHIN 30 SECONDS:

• CAN YOU THINK OF A WORD OF 1-6 SYLLABLES EACH?

Part 2

• CAN YOU HEAR ANY DIFFERENCE FROM THE TWO RECORDINGS?

• ROUTINE

WHICH ONE IS BETTER PRONOUNCED? WHY?

Read the words below, put the strongest syllable in the right place.

party   kindergarten   electricity   desert -dessert

Listen to the audios and try to imitate the quieter and stronger parts in each word.

PRACTICE

Listen to the audio and imitate the words

- Present
- Spectacular
- Photographer
- Supplement

MATERIALS TO BE RECORDED

How do you pronounce the words?
Listen to the attached native speaker's recordings and imitate the pronunciation.

- Satisfactory
- Analysis
- Inconclusive
- Harmonious
- Chinese
- Confidence

Module 2: Strength on sentence level

LESSON 2

STRONG AND QUIET WORDS

TASK A

Listen to the conversation and fill in the missing words.

- Why _____ you _____ such _____ hurry?
-I _____ run to join my Spanish class.
-What time _____ it start?
-_____ 10 minutes!
-Yeah, it is tight! Think about _____traffic. It's always heavy _____rush hours.
-You bet! Alright! I need to go ! Talk __ you later.

TASK B

Listen to the conversation again and fill in the missing words.

- _____ are you in such a _____?
-I gotta _____ to _____ my _____ class.
-What _____ does it _____?
-in 10 _____!
-Yeah, it's _____! Think about the _____. It's always heavy during ___ hours.
-You bet! Alright! I need to _____! _____ to you later.

THINK

Which task is harder to listen for the words? Why?

LISTEN

Listen to the recording and type the sentence.

Listen to this recording. Is it easier to hear all the words?

What is the difference between the two recordings?

RHYTHM OF ENGLISH

Read the phrases aloud. Then listen to the audios.
Is your reading the same as a native speaker's?

A bag of rice
A nice family
I can't!
She has a book.

Listen to the recordings again: which words are loud, which are quiet?

PRACTICE 1

1. Underline the strongest words in each sentence.

- She saw me yesterday.
- Could you check my homework?
- I made my coffee at home
- It's easy with good habits.
- His book was lost in the mall.

2. Check with the audios

THINK

Which words should be strong?

PRACTICE 2

Read the sentences and clap your hands on the strong words.

- Dogs chase cats.
- The dogs chase cats.
- The dogs chase the cats.
- The dogs will chase the cats.
- The dogs will be chasing the cats.

MATERIALS TO BE RECORDED

Listen to the audio and imitate the pronunciation. Send back your recordings.

1. I'll come tomorrow.
2. She's carrying the wrong bag.
3. The children will be swimming in the river.
4. I want to buy a car.
5. I hope to visit my grandparents.
6. I went out for shopping.

Module 3: Group the information

LESSON 3

GROUP THE INFORMATION

WARM UP

Can you add more information and make the sentences longer?

- He is the teacher that…

- I bought a book…

LISTEN AND BREAK INTO UNITS

1. She'd left when she had a baby and then decided not to go back although the job had been kept open for her

2. Tom dear where's the advert for this calculator because I don't know the address and I don't know who I've got to make the cheque payable to

LISTEN AND IMITATE

Repeat after the audios and pause when necessary.

- She is a teacher.
- A beautiful woman and her dog entered the room.
- When I woke up I didn't even realize what time it was.
- Because he was ill I didn't expect him to come to work.
- When I was in the canteen I tripped over a tray but my friend caught me.

Listen again and find the strongest words in each unit.

- She is a teacher.
- A beautiful woman and her dog/ entered the room.
- When I woke up/ I didn't even realize/ what time it was.
- Because he was ill/ I didn't expect him/ to come to work.
- When I was in the canteen/ I tripped over a tray/ but my friend caught me.

160

LISTEN AGAIN AND TRY TO IMITATE

- She is a **tea**cher.

- A <u>beautiful woman</u> and her **dog**/ <u>entered</u> the **room.**

- When I **woke up**/ I didn't even **realize**/ what **time** it was.

- Because he was **ill**/ I didn't **expect** him/ to come to **work.**

- When I was in the **canteen**/ I <u>tripped</u> over a **tray**/ but my <u>friend</u> **caught** me.

Module 4: talk about new information

LESSON 4

ADD RHYTHM INTO CONVERSATIONS

READ OUT CONVERSATIONS IN PAIRS

1. A: What color is your car?

   B: It is blue, dark blue.

2. A: Where are you going?

   B: To the supermarket. Where are you going?

   A: I am also going to the supermarket.

LISTEN TO THE RHYTHM

1. A: What color is your car?

   B: It is blue, dark blue.

2. A: Where are you going?

   B: To the supermarket. Where are **you** going?

   A: I am also going to the supermarket.

- Which words are stressed? Which are not? Why?
- Now you try to imitate.

MATCH

Listen to the audios and match each question with its answer.

1. a) Have you ever been to Spain before?   i)
   b) Did you work in the US?   ii)

2. a) That mobile looks familiar.   i)
   b) What is making that noise?   ii)

3. a) Who is talking on the phone?   i)
   b) What is Lucy doing?   ii)

4. a) She works at St. Mary's, doesn't she?   i)
   b) She works as a receptionist nearby, doesn't she?   ii)

I worked in Spain.

It's your phone.

Lucy is making a call.

She is the receptionist at the hospital.

PRACTICE

- Where is she from?

-She is from England. Where are you from?

-I am from China.

-How do you go to your office?

-I drive to the office.

-How long does it take?

-It's a 10-minute drive.

Now check the audios. Does it sound like a native speaker?

MATERIALS TO BE RECORDED

Listen to the audios and imitate the pronunciation. Send back your recordings.

1. -Which bag do you want to buy?   -The yellow bag.
2. -Do you want to go to Beijing?   - Yes, it's my dream city.
3. -How do you go to school?   -I walk to school.
4. -Where is your friend from?   -She is from California.
5. -How old is she?   -She is 9 years old.
6. -A: I've lost my hat.
    -B: What kind of hat?
    -A: It was a sun hat.

Module 5: contrastive information

LESSON 5

CONTRASTIVE INFORAMTION

Listen to the sentence said by a teacher in different ways, choose the correct meaning.

I wanted him to say the sentence.

Not write it down   Not her   Me, not you

1   2   3

And he did it   not the word

4   5

PRACTICE 1

Task A: Listen to the recordings. The answer 'I've always been terrified of spiders' is said in a slightly different way. Guess which question each answer matches to.

e.g. - Why don't you try keeping a spider as a pet?

      - I've always been terrified of spiders.

1. - Is there anything that really terrifies you?

   -

2. - How long have you been frightened of spiders?   A

   -

3. – Who's always been frightened of spiders?   B

   -   C

Task B: in pairs, read the questions and answers.

PRACTICE 2

Listen to the questions, and then read the answers in different ways (with proper strength on certain words). Then listen to the native recordings and compare with your reading.
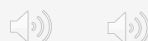
-Who bought a computer yesterday?

-Jane bought a computer yesterday.

-How did Jane get the computer yesterday?

-Jane bought the computer yesterday.

-What did Jane buy yesterday?

-Jane bought a computer yesterday.

-When did Jane buy a computer?

-Jane bought a computer yesterday.

-Who did Jane buy the computer with?

-Jane bought the computer with her parents.

164
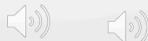
PRACTICE 3

Read the following sentences, with strength on the appropriate words, then listen to the recordings and compare.

1. A: Your brother's a doctor, isn't he?   B: No, my father's a doctor.
2. I think it's more important to have comfortable clothes than stylish ones.
3. A: I'll see you in the office on Friday.   B: I work at home on Fridays.
4. Watching football is much better than watching basketball.
5. Watching football is more interesting than playing football.

Sentence Stress Used at Its Best:
1 sentence, 7 different meanings!
By David Recine. www.WordStress.info

MATERIALS TO BE RECORDED

Listen to the audios and imitate the pronunciation. Send back your recordings.

1. It's my book, not yours.
2. Do you want to stay or leave?
3. I prefer that color.
4. Do you enjoy watching tennis or playing tennis?
5. Do you want to go to school or go home?
6. -Who cooked the dinner? -Your sister cooked it.

Module 6: Emphatic information

LESSON 6

EMPHATIC INFORMATION

Listen to the audios. Guess the implicit information conveyed by the speaker.

1. That was a difficult test.

2. This dish is so tasty.

3. I'm terribly sorry.

• Now listen to the audios and try to imitate.

- Read the sentences. Which word do you think is the strongest?

    1. My teacher went there by herself.
    2. The view is terrific.
    3. It's my own house.
    4. I spent a lot buying it.

- Listen to the audios. Check with your readings. Do they sound the same?

MORE EXAMPLES

- Some adverbs and modifiers are emphatic by nature.
- Can you think of any?

> (Hidden information)
> Indeed, utterly, absolute, terrific, tremendous, awful, always, great, really, truly, very, quite, entirely, extremely, only, own, alone, surely…

- Try to create scenarios and make sentences with them, read aloud your sentences.

MATERIALS TO BE RECORDED

Listen to the audios and imitate the pronunciation.

- 1. The watch is too expensive.
- 2. The food is awful.
- 3. It is extremely hot in summer.
- 4. It is my only chance.
- 5. He is surely fine.
- 6. She is always angry.

Module 7: Empty information

LESSON 7 EMPTY INFO

- Listen to the audios and write down the missing words.

  I must get my hair _____?

  She's got an essay _____.

  First I've got some shopping _____.

- Think. Why are these words quieter?

- Listen again and imitate.

- Can you come up with some more phrases like this?

PRACTICE 1

Read the sentences. Then listen to the audios to check your pronunciation.

- She's made a name for herself.
- We don't have any food left.
- I need to study hard from now.
- How is your work going?

PRACTICE 2

Read the sentences by yourself. Then listen to the audios to check your rhythm.

- She went out with Ken or somebody.
- Let me know if you need help or anything.
- Do you need water or something?
- I love hiking, cycling and these sorts of things.
- It lasts two hours or so.
- I'll pick you up at your place.

MATERIALS TO BE RECORDED

Listen to the audios and imitate the pronunciation.

- 1. Let's have the meeting at my place.
- 2. How's your father doing?
- 3. He lives along the coast or somewhere.
- 4. I don't know where to buy the stuff.
- 5. I have to arrive at 3 o'clock.
- 6. I've got some water to drink.

Module 8 is a final review of the presented content.

169

Appendix D. Qualitative questionnaire for Chinese participants

(For both explicit and implicit groups)

1. Please indicate your name, age, sex, and English level?

   a) Name:

   b) Age:

   c) Sex:

   d) Level of English:

2. Have you ever lived in an English-speaking country? If yes, how long?

3. Have you received a pronunciation training course before? If yes, what aspects did you deal with?

3 How important do you think pronunciation is for English learning?

   Not important at all   Very important

       1   2   3   4   5

4 How did you feel about your pronunciation before the course?

   Not good at all   Very good

       1   2   3   4   5

5 How useful do you think the training course is for improving your pronunciation?

   Not useful at all   Very useful

       1   2   3   4   5

6 To what extent do you think accurate pronunciation is necessary to communicate with native speakers?

   Not necessary at all   Very necessary

       1   2   3   4   5

7 How easy do you feel it is to follow the course?

   Not easy at all   Very easy

       1   2   3   4   5

8 If it is not easy for you, can you name the difficulties you have encountered?

9 How useful do you think our training course is for improving your pronunciation?

Very difficult   Not difficult at all

  1   2   3   4   5

10 Would you like to recommend this course to more students?

Not willing   very willing

  1   2   3   4   5

(Exclusively for explicit group)

11 How much can you understand the pronunciation rules taught in the course?

None of them   All of them

  1   2   3   4   5

12 Are you willing to continue improving other aspects of your English

pronunciation in this way?

Not willing   very willing

  1   2   3   4   5

13 Would you like native speakers' recordings to imitate for self-practice?

Not willing   very willing

  1   2   3   4   5

14 In which way do you feel the course has helped you?

15 Do you have any suggestion to help us improve the course?

(Exclusively for implicit group)

11 How much have you followed the audios and done the imitation practice?

0%   100%

1   2   3   4   5

12 How did you feel about doing the imitation practice?

Very difficult   Not difficult at all

1   2   3   4   5

13 To what extent do you think this imitation practice can improve your

pronunciation?

Not at all   a lot

1   2   3   4   5

14   Are you willing to continue the imitation practice in the future?

Not willing   very willing

1   2   3   4   5

16 Are you willing to continue improving other aspects of your English pronunciation

through imitation?

Not willing   very willing

1   2   3   4   5

17 In which way do you feel the course has helped you?

18 Do you have any suggestion to help us improve the course?

Appendix E. PRAAT script

Created by Leonardo Oliviera

```
form Extraction_from_TextGrids

    comment Directory where the binary Praat files are:

    sentence Folder C:\Path\

    comment Tier which contains annotation:

    integer Tier 2

    comment What measurements should be reported:

    boolean Duration 1

    boolean Intensity 1

    boolean Fundamental_frequency 1

    comment Report folder and name:

    sentence Output_file C:\Path\Results.txt

endform


#Create list of all files in folder

list = Create Strings as file list: "fileList", folder$ + "*.*"

files = Get number of strings

if files < 1

    exitScript: "No files at ", folder$ "."

endif


# Initialize variables

segDur =            undefined

segInt =           undefined

seg_Fzero =        undefined

seg_Fzero_max = undefined
```

```
# Write header of output file

writeFileLine: output_file$, "File", tab$, "Label",

... tab$, "Duration (s)", tab$, "Intensity (dB)", tab$, "Mean F0 (Hz)"


# Process all files in file list


for indxfile to files

    selectObject: list

    file$ = Get string: indxfile

    Read from file: folder$ + file$

    #selectObject:



    list_of_intervals# = Extract non-empty intervals... tier 0


    for i to size(list_of_intervals#)

        segment_Under_analysis = list_of_intervals#[i]

        selectObject: segment_Under_analysis

        name_seg$ = selected$ ("Sound")


        if duration == 1

            segDur = Get total duration

        endif


        if intensity == 1

            if segDur < 0.064

                segInt = undefined

            else

                selectObject: segment_Under_analysis
```

174

```
            To Intensity: 100, 0, "yes"

            segInt = Get maximum: 0, 0, "parabolic"

        endif

    endif


    if fundamental_frequency == 1

        selectObject: segment_Under_analysis

        To Pitch... 0.0 75 300

        seg_Fzero = Get mean: 0.0, 0.0, Hertz

        seg_Fzero_max = Get maximum: 0, 0, "Hertz", "parabolic"

    endif


    appendFileLine: output_file$, file$, tab$, name_seg$, tab$,
    ... segDur, tab$, segInt, tab$, seg_Fzero


  # In the future: add cleanup here. More than 10k objects crash Praat


    endfor
  endfor
```