

**Improving Intelligibility and Comprehensibility of  
Segmental and Suprasegmental Speech Patterns among  
Saudi Beginning-Level EFL Learners**

Turki Nafea Alharbi

A thesis submitted for the degree of Doctor of Philosophy of The Australian National  
University



November 2021

© Copyright by Turki Nafea Alharbi, 2021  
All Rights Reserved

## **Declaration**

I declare that, to the best of my knowledge, this thesis is my own original work and does not contain any material previously published or written by another person, except where otherwise indicated.

**Turki Nafea Alharbi**

**November 2021**

A handwritten signature in blue ink, appearing to read 'Turki', with a large, sweeping flourish extending to the right.

## **Acknowledgments**

First of all, I would like to show my profound gratitude to Dr Manuel Delicado Cantero for guiding and supervising this thesis. He was always available to provide constructive feedback, insightful suggestions, and support throughout my candidature. Special thanks go to Dr Ksenia Gnevsheva, who took me to the world of R, starting with her Quantitative Methods in Linguistics class. I would also thank Associate Professor Shunichi Ishihara for his valuable feedback at different stages of this process, especially in the write-up.

My heartfelt thanks to my family for all of their love and patience, especially during the pandemic. To my mother, Wadha, who still reminds me that I have promised her since 2013 that I will come back home. To my father, who passed away almost nine years ago but who remains with me always in spirit.

My sincere gratitude to Saudi Electronic University (SEU) for granting me a scholarship to pursue my doctoral degree. I must also thank Dr Rasha Alshaye (the head of the English Language and Translation Department at SEU) for her efforts during data collection. In addition, I would like to thank the Saudi Arabian Cultural Mission for making my studies much easier. Special thanks go to Ms Mashail Alkadi (the president of the Saudi Student Association in 2018-2019) for representing Saudi students in Australia through cultural, social, and intellectual initiatives for an Australian and non-Saudi audience, in particular the Cultural Exchange Salon. Such initiatives helped me find another life outside of academia in Canberra.

I would like to thank my colleagues in the Linguistics Program at ANU (Aser, Charbel, Gan, Maho, Mona, and Ruri) for being helpful, supportive, and wonderful officemates. The memories we made together will never be erased. My sincere appreciation goes to the Higher Degree Research administration (especially Rebecca) in the School of Languages, Linguistics and Literature for minimising administrative challenges during my candidature. I would also like to thank Rozana Muir in Student Central for being kind, supportive, and welcoming. In

addition, I would like to thank Nathanael Shook for editing this thesis. Furthermore, I would like to thank the external examiners for their informative feedback. Finally, I must thank the participants and raters for their assistance, without which this work would not have been possible.

## **Dedication**

To her. When you believe in me, I believe every dream is achievable.

2 November 2021  
Canberra

## Abstract

In the last two decades, second language (L2) pronunciation education and research have emphasised the need for intelligibility and comprehensibility over the traditional goal of achieving native-like fluency. However, it remains unclear what instructors should focus on to promote intelligibility and comprehensibility (e.g., segmentals vs. suprasegmentals) among beginning-level learners in the Saudi EFL context. Another uncertainty is how long their pronunciation gains might last after instruction. Although motivation has been significantly related to L2 pronunciation development, it is unclear how this factor is related to developing speech intelligibility and comprehensibility among beginning-level EFL learners. Learners' perspectives about pronunciation instruction have also been largely neglected, and studies have not elicited in-depth comments from native English-speaking raters about these learners' difficulties.

The present study recruited 45 participants from a beginning-level EFL classroom at a Saudi university to complete a battery of pre-, immediate post-, and delayed post-tests. Participants were divided into three groups. The first experimental group (the segmental group) of 15 participants received explicit pronunciation instruction on segmentals. The second experimental group (the suprasegmental group) of 15 participants received explicit pronunciation instruction on suprasegmentals. The control group of 15 participants was given non-explicit pronunciation instruction.

A series of mixed-effects models was fit in R to predict significant differences within and between the groups across tests. The segmental group's results for the immediate post-test revealed that intelligibility and comprehensibility improved significantly over the pre-test and control group, with the exception of the segment /tʃ/. However, the group failed to attain intelligibility and comprehensibility improvement over time, except for segment /v/. Regarding the suprasegmental group, the immediate post-test showed significant improvement in

intelligibility and comprehensibility over the pre-test and control group. In addition, the suprasegmental group showed improvement over time, with the exception of rhythmic patterns in terms of intelligibility and stress in terms of comprehensibility. The suprasegmental group was more motivated than the segmental and control groups in the “intended efforts” and “ideal L2 self” dimensions, while participants in the control group were more motivated to learn English than the two experimental groups in the “ought-to L2 self” dimension. In addition, the suprasegmental group showed more positive attitudes toward learning English than the segmental and control groups. Finally, the experimental groups found the intervention helpful, despite it being relatively short.

Taken together, this study contributes to the L2 pronunciation field by demonstrating that pronunciation instruction could be successfully incorporated into a beginning-level Saudi EFL classroom, according to the segmental and suprasegmental results. The delayed post-test results for the suprasegmental group suggested the effect of pronunciation instruction could be sustained over time. In addition, the findings showed L2 motivation to be significantly related to greater improvement in pronunciation. Based on these empirical results, the study provides a pedagogical description of how to improve pronunciation teaching for beginning-level learners in a Saudi EFL context.

## Table of Contents

Declaration .....	ii
Acknowledgments .....	iii
Dedication .....	v
Abstract .....	vi
List of Tables.....	xiii
List of Figures .....	xv
Chapter 1: Introduction .....	1
1.1 Overview and Statement of the Problem.....	1
1.2 Research Objectives .....	4
1.3 Significance of the Study.....	6
1.4 Thesis Structure .....	9
Chapter 2: EFL Instruction in Saudi Arabia.....	11
2.1 Introduction .....	11
2.2 The Saudi Education System and EFL Instruction.....	12
2.3 Methods, Techniques, and Materials for Teaching EFL .....	16
2.4 Essential Reforms for Teaching EFL in the Saudi Context .....	20
2.5 Chapter Summary .....	22
Chapter 3: Literature Review .....	23
3.1 Introduction .....	23
3.2 Theoretical Framework of L2 Phonology .....	23
3.2.1 SLM .....	24
3.2.2 PAM and PAM-L2 .....	27
3.2.3 Comparisons between SLM and PAM/-L2 .....	31
3.3 Segmental and Suprasegmental Aspects of English Pronunciation .....	32
3.3.1 English Segmentals.....	32
3.3.1.1 Classification of English Segmentals .....	32
3.3.1.2 Common English Segmental Errors among Saudi EFL Learners .....	34
3.3.2 English Suprasegmentals .....	34
3.3.2.1 Lexical Stress .....	35
3.3.2.2 Rhythm.....	36
3.3.2.3 Final Intonation.....	38
3.4 Intelligibility and Comprehensibility in L2 Pronunciation .....	39
3.5 The Efficacy of L2 Pronunciation Instruction.....	41

3.5.1	Intelligibility and Nativeness Principles .....	41
3.5.2	Effect of Instruction on Segmentals and Suprasegmentals.....	42
3.5.3	Effect of L2 Pronunciation Instruction over Time .....	47
3.5.4	Effect of L2 Pronunciation Instruction on Beginning-Level Learners.....	48
3.6	Relationship of L2 Motivation to Intelligible and Comprehensible Speech .....	50
3.7	Learner Perceptions of Pronunciation Instruction Effectiveness .....	56
3.8	Native-Speaker Perceptions of Non-Native Speakers .....	57
3.9	Research Questions .....	59
3.10	Chapter Summary .....	60
Chapter 4:	Methodology.....	61
4.1	Introduction .....	61
4.2	Research Design .....	61
4.3	Participants and Setting .....	64
4.4	Instruments .....	66
4.4.1	Questionnaires .....	67
4.4.2	Segmental Stimuli.....	69
4.4.3	Suprasegmental Stimuli .....	72
4.5	Pronunciation Intervention .....	75
4.5.1	Pronunciation Intervention Framework.....	75
4.5.2	Pronunciation Intervention Materials .....	78
4.6	Data Collection Procedures .....	79
4.6.1	Phase I: Recruitment, Background, Pretest, and Motivation Questionnaire .....	80
4.6.2	Phase II: Teacher Preparation, Intervention, Immediate Posttest, and Post- Intervention Questionnaire .....	81
4.6.3	Phase III: Delayed Posttest .....	84
4.6.4	Phase IV: Perceptual Assessment of Recordings and Post-Assessment Questionnaire .....	84
4.7	Operationalisation of Speech Intelligibility and Comprehensibility Constructs .....	87
4.8	Raters and Inter-Rater Reliability of the Audio-Recording Assessments .....	89
4.9	Pilot Testing.....	90
4.10	Validity of the Study .....	92
4.11	Statistical Analysis Procedures.....	94
4.12	Chapter Summary .....	98
Chapter 5:	Results .....	99
5.1	Introduction .....	99

5.2	Intelligibility of Segmentals .....	99
5.2.1	The Segment /p/ .....	99
5.2.1.1	Descriptive Results for /p/ Intelligibility .....	99
5.2.1.2	Modelling Results for /p/ Intelligibility .....	100
5.2.2	The Segment /v/ .....	103
5.2.2.1	Descriptive Results for /v/ Intelligibility .....	103
5.2.2.2	Modelling Results for /v/ Intelligibility .....	104
5.2.3	The Segment /ʃ/ .....	106
5.2.3.1	Descriptive Results for /ʃ/ Intelligibility .....	106
5.2.3.2	Modelling Results for /ʃ/ Intelligibility .....	107
5.2.4	The Segment /ŋ/ .....	109
5.2.4.1	Descriptive Results for /ŋ/ Intelligibility .....	109
5.2.4.2	Modelling Results for /ŋ/ Intelligibility .....	110
5.3	Comprehensibility of Segmentals.....	112
5.3.1	Descriptive Results .....	112
5.3.2	Modelling Results for Comprehensibility of Segmentals.....	114
5.4	Intelligibility of Suprasegmentals.....	116
5.4.1	Lexical Stress.....	116
5.4.1.1	Descriptive Results for Intelligibility of Lexical Stress.....	116
5.4.1.2	Modelling Results for Intelligibility of Lexical Stress .....	117
5.4.2	Rhythm .....	119
5.4.2.1	Descriptive Results for Intelligibility of Rhythm .....	119
5.4.2.2	Modelling Results for Intelligibility of Rhythm .....	120
5.4.3	Final Intonation.....	122
5.4.3.1	Descriptive Results for Intelligibility of Final Intonation .....	122
5.4.3.2	Modelling Results for Intelligibility of Final Intonation .....	123
5.5	Comprehensibility of Suprasegmentals .....	125
5.5.1	Lexical Stress.....	125
5.5.1.1	Descriptive Results for Comprehensibility of Lexical Stress .....	125
5.5.1.2	Modelling Results for Comprehensibility of Lexical Stress.....	126
5.5.2	Rhythm .....	127
5.5.2.1	Descriptive Results for Comprehensibility of Rhythm.....	127
5.5.2.2	Modelling Results for Comprehensibility of Rhythm .....	129
5.5.3	Final Intonation.....	130

5.5.3.1 Descriptive Results for Comprehensibility of Final Intonation.....	130
5.5.3.2 Modelling Results for Comprehensibility of Final Intonation.....	131
5.6 L2 Motivation as Related to Pronunciation Improvements.....	133
5.6.1 Descriptive Results .....	133
5.6.2 Modelling Results for Motivation .....	135
5.7 How Learners Perceived the Treatment .....	137
5.7.1 Descriptive Results of the Closed-Ended Questions .....	137
5.7.2 Modelling Results of the Closed-Ended Questions.....	138
5.7.3 Narrative Responses for the Open-Ended Questions.....	139
5.8 Raters' Perceptions of the Pronunciation of Participants .....	140
5.8.1 Perceived Intelligibility of Segmental and Suprasegmental Features .....	140
5.8.2 Perceived Comprehensibility of Segmental and Suprasegmental Features.....	142
5.9 Chapter Summary .....	143
Chapter 6: Discussion.....	144
6.1 Introduction .....	144
6.2 Summary of Major Findings .....	144
6.3 Research Question 1: Immediate Effect on Segmentals.....	147
6.4 Research Question 2: Immediate Effect on Suprasegmentals.....	149
6.5 Research Question 3: Long-Term Effectiveness of Treatment .....	151
6.6 Research Question 4: Motivations and Attitudes as Predictors.....	155
6.7 Research Question 5: Learners' Perceptions of the Treatment .....	158
6.8 Research Question 6: Raters' Perceptions of Learners .....	159
6.9 Chapter Summary .....	161
Chapter 7: Implications for L2 Pronunciation Pedagogy and Research .....	162
7.1 Introduction .....	162
7.2 Implications for L2 Pronunciation Pedagogy.....	162
7.2.1 How Can Segmental and Suprasegmental Pronunciation Be Taught in the Saudi Beginning-Level EFL Classroom? .....	163
7.2.2 How Can Pronunciation Gains Be Attained over Time in the Saudi Beginning-Level EFL Classroom? .....	166
7.2.3 How Can Motivation Be Addressed in the Saudi Beginning-Level EFL Classroom? .....	167
7.2.4 What Do Learner Perceptions Suggest about the Effectiveness of L2 Pronunciation Instruction in the Saudi Beginning-Level EFL Classroom? .....	168
7.3 Implications for L2 Pronunciation Research.....	169
7.4 Chapter Summary.....	171

Chapter 8: Conclusion .....	172
8.1 Overview of the Study .....	172
8.2 Original Contributions of the Study .....	172
8.3 Limitations and Recommendations for Future Research .....	174
References .....	178
Appendix A: Demographic and Language Background Questionnaire (English Version) ...	191
Appendix B: L2 Motivation Questionnaire (English Version) .....	195
Appendix C: Post-Intervention Questionnaire (English Version).....	199
Appendix D: Questionnaire for Native English Raters .....	201
Appendix E: Segmental Stimuli .....	204
Appendix F: Suprasegmental Stimuli .....	206
Appendix G: Approval Letter to Conduct the Study.....	208
Appendix H: Participant Information Sheet for Students (English Version).....	209
Appendix I: Participant Information Sheet for Students (Arabic Version).....	214
Appendix J: Written Consent for Students (English Version) .....	218
Appendix K: Written Consent for Students (Arabic Version) .....	219
Appendix L: Demographic and Language Background Questionnaire (Arabic Version) .....	220
Appendix M: L2 Motivation Questionnaire (Arabic Version).....	223
Appendix N: Post-Intervention Questionnaire (Arabic Version).....	226
Appendix O: Participate in a Paid Research Project .....	228
Appendix P: Audio-Recordings Assessment Instructions .....	229
Appendix Q: Participant Information Sheet for Raters .....	231
Appendix R: Written Consent for Raters .....	235

## List of Tables

Table 3.1: English Phonemic Consonant Inventory .....	33
Table 3.2: English Consonants without a Northern Saudi Arabic Equivalent and Common Substitutions .....	34
Table 3.3: Possible Intelligibility and Comprehensibility Combinations .....	40
Table 4.1: Participants' Language Background Information .....	65
Table 4.2: Goals of Each Stage of Development .....	76
Table 4.3: Using Modality to Teach Stress Patterns .....	77
Table 4.4: Textbook Summary .....	79
Table 4.5: Data Collection Phases.....	80
Table 4.6: Division of Audio-Recordings into Blocks.....	86
Table 4.7: Operationalisation of Constructs.....	88
Table 4.8: Modelling Analysis Procedures for Segmental and Suprasegmental Features.....	97
Table 4.9: Modelling Analysis Procedures for Motivation and Post-Intervention Questionnaire .....	97
Table 5.1: Descriptive Statistical Summary for /p/ Intelligibility .....	99
Table 5.1: Model for /p/ Intelligibility with Pretest as Reference Level.....	101
Table 5.3: Model for /p/ Intelligibility with Posttest as Reference Level .....	102
Table 5.4: Descriptive Statistical Summary for /v/ Intelligibility .....	103
Table 5.5: Model for /v/ Intelligibility with Pretest as Reference Level.....	105
Table 5.6: Model for /v/ Intelligibility with Posttest as Reference Level .....	105
Table 5.7: Descriptive Statistical Summary for /tʃ/ Intelligibility .....	106
Table 5.8: Model for /tʃ/ Intelligibility with Pretest as Reference Level.....	108
Table 5.9: Model for /tʃ/ Intelligibility with Posttest as Reference Level .....	108
Table 5.10: Descriptive Statistical Summary for /ŋ/ Intelligibility .....	109
Table 5.11: Model for /ŋ/ Intelligibility with Pretest as Reference Level.....	111
Table 5.12: Model for /ŋ/ Intelligibility with Posttest as Reference Level .....	112
Table 5.13: Descriptive Statistical Summary for Comprehensibility of Segmentals.....	113
Table 5.14: Model for Comprehensibility of Segmentals with Pretest as Reference Level ..	114
Table 5.15: Model for Comprehensibility of Segmentals with Posttest as Reference Level.	115
Table 5.16: Descriptive Statistical Summary for Intelligibility of Lexical Stress .....	116
Table 5.17: Model for Intelligibility of Lexical Stress with Pretest as Reference Level.....	117
Table 5.18: Model for Intelligibility of Lexical Stress with Posttest as Reference Level .....	118
Table 5.19: Descriptive Statistical Summary for Intelligibility of Rhythm.....	119
Table 5.20: Model for Intelligibility of Rhythm with Pretest as Reference Level.....	120

Table 5.21: Model for Intelligibility of Rhythm with Posttest as Reference Level .....	121
Table 5.22: Descriptive Statistical Summary for Intelligibility of Final Intonation .....	122
Table 5.23: Model for Intelligibility of Final Intonation with Pretest as Reference Level....	123
Table 5.24: Model for Intelligibility of Final Intonation with Posttest as Reference Level ..	124
Table 5.25: Descriptive Statistical Summary for the Comprehensibility of Lexical Stress...	125
Table 5.26: Model for Comprehensibility of Lexical Stress with Pretest as Reference Level.....	126
Table 5.27: Model for Comprehensibility of Lexical Stress with Posttest as Reference Level.....	127
Table 5.28: Descriptive Statistical Summary for Comprehensibility of Rhythm .....	128
Table 5.29: Model for Comprehensibility of Rhythm with Pretest as Reference Level.....	129
Table 5.30: Model for Comprehensibility of Rhythm with Posttest as Reference Level .....	130
Table 5.31: Descriptive Statistical Summary for Comprehensibility of Final Intonation .....	131
Table 5.32: Model for Comprehensibility of Final Intonation with Pretest as Reference Level.....	132
Table 5.33: Model for Comprehensibility of Final Intonation with Posttest as Reference Level.....	132
Table 5.34: Descriptive Statistical Summary for Motivational Dimensions .....	134
Table 5.35: Model for Motivation.....	135
Table 5.36: Descriptive Statistical Summary for Post-Intervention Theme Score .....	137
Table 5.37: Model for the Theme of the Post-Intervention Questionnaire .....	139
Table 6.1: A Summary of Segmental and Suprasegmental Results across a Battery of Tests .....	145

## List of Figures

Figure 3.1. Outline of PAM/-L2 patterns and predictions based on L1-L2 similarities (adopted from Almbark, 2012, p. 47).....	30
Figure 5.1. Likelihood of /p/ intelligibility. ....	100
Figure 5.2. Model prediction of /p/ intelligibility based on time (pretest, posttests), group (experimental, control), and position (onset, coda).....	103
Figure 5.3. Likelihood of /v/ intelligibility. ....	104
Figure 5.4. Model prediction of /v/ intelligibility based on time (pretest, posttests) and group (experimental, control). ....	106
Figure 5.5. Likelihood of /tʃ/ intelligibility in pretest and posttests. ....	107
Figure 5.6. Model prediction of /tʃ/ intelligibility based on time (pretest, posttests) and group (experimental, control). ....	109
Figure 5.7. Likelihood of /ŋ/ intelligibility in pretest and posttests. ....	110
Figure 5.8. Model prediction for /ŋ/ intelligibility based on time (pretest, posttests) and group (experimental, control). ....	112
Figure 5.9. Mean comprehensibility scores for segmentals in pretest and posttests.....	113
Figure 5.10. Model prediction for the comprehensibility of segmentals based on time (pretest, posttests) and group (experimental, control). ....	115
Figure 5.11. Likelihood of intelligibility of lexical stress.....	117
Figure 5.12. Model prediction for intelligibility of lexical stress based on time (pretest, posttests) and group (experimental, control). ....	118
Figure 5.13. Mean intelligibility scores for rhythm. ....	119
Figure 5.14. Model prediction for intelligibility of rhythm based on time (pretest, posttests) and group (experimental, control). ....	121
Figure 5.15. Likelihood of intelligibility of final intonation.....	122
Figure 5.16. Model prediction for intelligibility of final intonation based on time (pretest, posttests) and group (experimental, control). ....	124
Figure 5.17. Mean comprehensibility scores of lexical stress.....	125
Figure 5.18. Model prediction for comprehensibility of lexical stress based on time (pretest, posttests) and group (experimental, control). ....	127
Figure 5.19. Mean comprehensibility scores for rhythm. ....	128
Figure 5.20. Model prediction for comprehensibility of rhythm based on time (pretest, posttests) and group (experimental, control). ....	130
Figure 5.21. Mean comprehensibility scores of final intonation.....	131
Figure 5.22. Model prediction for comprehensibility of final intonation based on time (pretest, posttests) and group (experimental, control). ....	133
Figure 5.23. Mean scores across the four motivational dimensions. ....	134
Figure 5.24. Model prediction for motivation score in the interaction between motivational	

dimensions and groups. ....	136
Figure 5.25. Post-intervention mean scores for satisfaction and outcomes. ....	138
Figure 5.26. Model prediction for the interaction between theme (intervention satisfaction, outcomes) and group (segmental, suprasegmental). ....	139

## **Chapter 1: Introduction**

### **1.1 Overview and Statement of the Problem**

A long-standing debate in the literature on second/foreign language acquisition (SLA/FLA) is whether adult second language (L2) learners can improve their pronunciation through explicit instruction (Derwing & Munro, 2015; Thomson & Derwing, 2015). To make L2 pronunciation instruction more effective, research has suggested that L2 learners should not be expected to attain fully native-like performance, which appears beyond the goal of attaining intelligible and comprehensible communication between speakers (Jenkins, 2000; Levis, 2005, 2018; Munro & Derwing, 1995). In addition, Munro and Derwing (2015) suggested that successful L2 spoken communication is the responsibility of non-native speakers and native listeners alike (p. 388). In other words, native listeners of an L2 need to put effort into maintaining spoken communication with non-native speakers.

On the other hand, some researchers have pointed out that late/adult L2 learners could attain native-like performance, particularly if learners are exposed to a large amount of L2 input or if the first language (L1) and L2 are related (e.g., Bongaerts et al., 1997; Piller, 2002). Nevertheless, the main goal of teaching L2 pronunciation is to make learners more intelligible and comprehensible without setting the extremely difficult goal of native-like pronunciation. While some L2 learners might be able or want to sound exactly like a native speaker, it is important to bear in mind that pronunciation improvements do not necessarily mean “native-like”. To support this claim, Thomson and Derwing (2015) stated that “native-like pronunciation is an unrealistic goal, but that improved intelligibility and comprehensibility are achievable, and that L2 pronunciation research focusing on these speech dimensions is of more practical value than studies of accentedness alone” (p. 335). However, with this in mind, more

research is needed to explore whether explicitly teaching English pronunciation to beginning-level learners is a realistic goal to promote intelligibility and comprehensibility.

As a result, intelligibility and comprehensibility have become core concepts in research on L2 pronunciation because the ultimate goal of human communication is to be understood (Jenkins, 2000; Levis, 2005). Intelligibility refers to the ability of speakers' utterances to be clearly understood, while comprehensibility deals with how difficult it is for listeners to understand speakers' utterances (Levis, 2005; Munro & Derwing, 1995; Munro et al., 2006). Furthermore, Jenkins (2013) indicated that the importance of L2 pronunciation teaching has shifted from achieving native-like performance to maintaining a satisfactory level of intelligibility and comprehensibility.<sup>1</sup> Several factors have led to this shift, such as the number of non-native speakers of English being higher than the number of native speakers, leading to greater emphasis on meaningful communication (Jenkins, 2000).

Nowhere is this need more apparent than in developing countries. Saudi Arabia, for example, has invested heavily in English as a foreign language (EFL) education to promote economic opportunities and development. Although English has no official status in Saudi Arabia and is considered a foreign language, it is often used in certain situations and academic institutions (Alrashidi & Phan, 2015).<sup>2</sup> In the past few years, the Ministry of Education—formerly the Ministry of Higher Education—has mandated that all public universities use English as a medium of instruction. As a consequence, the Intensive English Program (IEP) has been established in the Preparatory Year Program (PYP) that students take before entering university. IEP offers a range of English courses in English for Specific Purposes (ESP) or English for General Purposes (EGP). The purpose of this programme is to improve the English

---

<sup>1</sup> The value of intelligibility and comprehensibility in L2 pronunciation education is illustrated in Section 3.4.

<sup>2</sup> A detailed review of EFL education in Saudi Arabia is given in Chapter 2.

skills of university students, especially those who will have English-medium instruction (Ebad, 2014).

There are several problems with IEP in this context. One is that most Saudi EFL learners show only a basic level of proficiency. Within two semesters, they are expected to reach an advanced level to pass the programme. Although students in Saudi Arabia begin to learn English from elementary school, most do not become proficient in English (Alrabai, 2016). A second problem is that Saudi EFL learners are only exposed to English in the classroom. Without an English-speaking environment, they lack opportunities for L2 exposure outside the classroom (Alharbi, 2015; Al-Seghayer, 2015). Given these problems, it is essential to make English communication in the IEP classroom between Saudi EFL learners and their teachers more intelligible and comprehensible. Thus, teachers and students have to overcome the problems associated with not being in an English-speaking country and make the best use of classroom opportunities.

A third problem is that policymakers, EFL teachers, and IEP designers struggle to make the English class more beneficial and effective. For instance, due to a lack of studies on teaching IEP in Saudi Arabia, policymakers have recruited native English-speaking teachers (Liton, 2013; Ur Rahman & Alhaisoni, 2013). This decision seeks to minimise Arabic in the English class and increase the class's effectiveness. Unfortunately, Saudi EFL learners are not always sufficiently intelligible and comprehensible for their native-speaking teachers, especially early on, which could hamper learning. Therefore, policymakers need to know how to make English communication in the class more intelligible and comprehensible to improve learning outcomes.

A fourth problem is that it is unclear to what extent L2 motivation is related to improvements in the pronunciation of Saudi beginning-level EFL learners in terms of intelligible and comprehensible speech. A fifth problem is that it is unclear what these learners

would think about a pronunciation intervention. Therefore, it is important to examine these perspectives as well. A sixth problem is what native speakers think about pronunciation challenges that impede Saudi beginning-level EFL learners from being more intelligible and comprehensible.

Many researchers have emphasised the importance of teaching pronunciation to L2 learners (e.g., Couper, 2006; Derwing & Munro, 2015; Isaacs, 2009; Levis, 2005, 2018). However, the EFL classroom in Saudi Arabia remains understudied concerning this topic, resulting in several gaps in the literature. For instance, to what extent are Saudi beginning-level EFL learners intelligible and comprehensible in English? How can their intelligibility and comprehensibility be improved? What segmental and suprasegmental features affect their intelligibility and comprehensibility the most? What instructional approaches would best help them develop more intelligible and comprehensible pronunciation? Is L2 motivation related in this regard? What do learners think about the effectiveness of pronunciation instruction? Finally, what do native listeners think about the speech intelligibility and comprehensibility of Saudi beginning-level EFL learners?

To answer these questions, further work is needed on English pronunciation instruction in Saudi Arabia from a classroom perspective. Thus, this study investigated the influence of explicit pronunciation instruction on the intelligibility and comprehensibility of Saudi EFL learners, with a focus on implications for teaching to increase the effectiveness of the EFL class.<sup>3</sup>

## **1.2 Research Objectives**

The primary objective of this study was to explore whether explicit pronunciation instruction could be effective at improving Saudi beginning-level EFL learners' speech

---

<sup>3</sup> The effectiveness of L2 pronunciation instruction is reviewed in Section 3.5.

intelligibility and comprehensibility by targeting segmental and suprasegmental features. To date, it is unclear whether segmental or suprasegmental features cause more difficulties with the speech intelligibility and comprehensibility of these learners and what approach to address these difficulties would be more effective in the Saudi classroom.

The second objective was to determine whether these learners retained intelligibility and comprehensibility improvements over time. The third objective was to explore whether L2 motivation was related to intelligibility and comprehensibility improvements. L2 motivation has been found to be a vital factor to explain why some L2 learners are more intelligible and comprehensible than others (Moyer, 1999; Nagle, 2018; Piske et al., 2001; Saito et al., 2017). Emphasising the significance of individual variability in L2 speech learning (e.g., in terms of motivation), Thomson and Derwing (2015) stated that “complementary qualitative analyses should be conducted to provide insights into individual differences in learning, such as motivation, the nature of interactions in the L2, and other social influences. Such analyses can reveal evidence that quantitative research cannot access” (p. 328).

However, it is important to note that the primary aim of examining motivational orientations was to explore the relationship between group (segmental, suprasegmental, and control) and the four motivational profiles. These data were expected provide a better understanding of how non-pronunciation factors affected the L2 pronunciation development of learners who received explicit pronunciation instruction. Therefore, this thesis did not seek to explore causation but rather the relationship between L2 motivation and the development of L2 speech among Saudi beginning-level EFL learners.

The fourth objective was to explore the perceptions of Saudi beginning-level EFL learners about the effectiveness of pronunciation instruction. Such feedback could help enhance this type of instruction. The fifth aim was to investigate how native English speakers assessed

the intelligibility and comprehensibility of Saudi learners. Their perceptions could help determine what segmental and suprasegmental features impede pronunciation.

### **1.3 Significance of the Study**

This study bridges an important gap in the literature and is significant for several reasons. To the best of my knowledge, it is the first to examine a specific set of segmental and suprasegmental features with Saudi beginning-level EFL learners to help them achieve more intelligible and comprehensible speech. Several researchers, such as Zielinski and Yates (2014), have called for teaching pronunciation to beginning-level ESL learners to maximise their confidence in oral communication:

When pronunciation difficulties lead to a reticence or loss of confidence to interact in spoken English, beginning-level learners have limited opportunities for practice, which, in turn, can affect further language development. This lack of progress in L2 acquisition can ultimately contribute to social isolation and limit educational and work opportunities [...] It is, therefore, of utmost importance that pronunciation instruction be an integral part of the process of learning English from the beginning and not regarded as an option added at a later stage. (p. 59)

Following the same line of argument, Darcy (2018) stressed that “Since pronunciation improvement takes time, it is worth seriously considering whether beginners should be taught pronunciation when they are learning grammar and vocabulary” (p. 20). This is even more important for beginning-level EFL learners as they have fewer opportunities for exposure to English outside the classroom. Thus, it is a crucial stage to promote their intelligibility and comprehensibility because learners’ phonological competence can be greatly shifted in the first six months of massive exposure to an L2 (Derwing & Munro, 2015). This study was a reaction to that call from an EFL perspective. The study also offers pedagogical and practical implications to improve pronunciation in the Saudi EFL classroom early on.

Although L2 pronunciation has recently become a more central issue for teachers, learners, and researchers, it is still understudied in some areas, such as the Saudi EFL context.

For instance, most previous studies on the status of English pronunciation instruction have been conducted in an English-speaking environment, where English is dominant (e.g., Derwing & Munro, 2013; Kennedy & Trofimovich, 2010; Saito & Lyster, 2012; Saito et al., 2016). In contrast, this study examined pronunciation in an EFL context.

Furthermore, most previous studies in the Saudi EFL context examined pronunciation issues among advanced and intermediate learners (e.g., Ahmad & Muhiburrahman, 2013; Ahmad & Nazim, 2014; Al-Ahdal et al., 2015; Ali, 2015; Aljumah, 2011; Al-Wossabi, 2016; Hamad, 2013; Hameed & Aslam, 2015; Rajab, 2013; Shah & Al-Bargi, 2016). For this reason, the current study focused on beginning-level EFL learners. Furthermore, no previous research in a Saudi EFL setting had examined the intelligibility and comprehensibility constructs as ultimate goals of teaching pronunciation. This study addressed that gap by exploring whether beginning-level learners were likely to move toward intelligible and comprehensible speech.

Most previous investigations into how to improve L2 pronunciation have had a lab-based perspective with only segmental experiments (e.g., Bradlow et al., 1997; Iverson et al., 2005; Iverson et al., 2012; Kartushina et al., 2015; Lee & Lyster, 2017) or suprasegmental experiments (e.g., Chun et al., 2013; Field, 2005; Hardison, 2004; Hirata, 2004; Huensch & Tremblay, 2015). However, such studies might not be applicable to the classroom, outside of ideal settings. Thus, this study was conducted in the more naturalistic setting of an EFL classroom.

There continues to be much debate about the effectiveness of pronunciation strategies, activities, and techniques in the Saudi EFL classroom (e.g., Ahmad & Nazim, 2014; Alfallaj, 2013; Daif-Allah & Khan, 2016; Hameed & Aslam, 2015). Thus, the current study could help policymakers, teachers, and curriculum designers in Saudi Arabia better understand this setting

in order to determine the most effective pronunciation strategies, activities, and techniques for improving intelligibility and comprehensibility.

Saudi EFL learners have to overcome major pronunciation challenges in class to complete the PYP in one year. Therefore, this study could help these learners reach this goal by improving their communication. The ultimate goal of the pronunciation intervention was to explore whether learners maintained improvements over time (through a delayed posttest). Munro and Derwing (2015) suggested this type of test was optimal to explore whether L2 learners retained pronunciation gains over time (p. 84). In addition, Thomson and Derwing (2015) stated that “the ideal study should include a delayed posttest to determine whether the intervention had a lasting effect” (p. 327). Nevertheless, as far as I am aware, no previous research in the Saudi EFL context has implemented a delayed posttest phase. Thus, the present study implemented a delayed posttest to examine whether pronunciation gains lasted longer with Saudi beginning-level EFL learners.

Furthermore, there is a lack of research on how L2 motivation could be related to improvements in intelligibility and comprehensibility among Saudi beginning-level EFL learners, as a great deal of research has focused on motivation as related to achievement in a course or language based on the L2 Motivational Self System (e.g., Alshehri, 2009; Altalib, 2019; Moskovsky et al., 2016). Therefore, this study examined this relationship between L2 motivation and the development of L2 pronunciation.

There is also a lack of research exploring how to enhance the effectiveness of pronunciation instruction based on L2 learners’ perceptions, and previous research that has asked learners about this has largely been in an English-speaking country (e.g., Alameen, 2014; Derwing, 2003). Hence, this study explored what Saudi beginning-level learners thought about the effectiveness of pronunciation instruction.

In addition, it is vital to examine native speakers' in-depth comments about Saudi beginning-level EFL learners to determine which segmental and suprasegmental features are more challenging. Little research, however, has asked native-speaking raters about the pronunciation of non-native speakers (e.g., Gordon & Darcy, 2016; Zielinski, 2008). As a result, this study asked raters, after they completed the perceptual assessments, to comment on learners' intelligibility and comprehensibility.

#### **1.4 Thesis Structure**

This thesis consists of eight chapters. Chapter 1 presents an overview of the study, including the problems, objectives, and significance of the study. Chapter 2 gives an overview of the Saudi education system and EFL in each educational stage. Methods, techniques, and materials are also discussed, followed by a review of education reforms.

Chapter 3 offers a critical review of the L2 pronunciation literature, including a theoretical overview of phonology and suprasegmental aspects of English pronunciation. Intelligibility and comprehensibility are defined and operationalised as key concepts. In addition, a wide range of issues under L2 pronunciation instruction is reviewed. The relationship between L2 motivation and the development of intelligibility and comprehensibility is discussed, along with learners' perceptions about the effectiveness of L2 pronunciation instruction and the perceptions of native English speakers about the L2 pronunciation of non-native speakers. Finally, the research questions are stated.

Chapter 4 explains the study's methodology, beginning with the rationale for the research design and participant demographics. The chapter then describes the instruments, intervention framework, and materials. This is followed by a detailed description of the four phases of data collection. After that, the chapter explains the operationalisation of intelligibility and comprehensibility of segmental and suprasegmental features and explains the raters and

inter-rater reliability. To improve the main study, a pilot test was conducted. The chapter explains how the study controlled risks to internal and external validity and how the data were analysed.

Chapter 5 reports the findings, starting with the mixed-effects models of segmental and suprasegmental features for intelligibility and comprehensibility in order to answer the first three research questions. The results of the mixed-effects models from the L2 motivation questionnaire are presented to answer Research Question 4, learners' perceptions about the explicit intervention are reported to answer Research Question 5, and the raters' perceptions about learners' pronunciation are presented to answer Research Question 6.

Chapter 6 discusses the findings in relation to the research questions. After a summary of the major findings, the chapter discusses the first three research questions based on the immediate and delayed posttests. This is followed by a discussion of the other three research questions. Chapter 7 describes practical, pedagogical, and research implications for pronunciation in the Saudi beginning-level EFL classroom. These implications could help academics, instructors, policymakers, curriculum designers, and researchers make Saudi EFL classes more successful. Chapter 8 concludes the study with its contributions, limitations, and recommendations for future research.

## **Chapter 2: EFL Instruction in Saudi Arabia**

### **2.1 Introduction**

Despite the growing prominence of EFL in the Saudi education system, pronunciation is often little more than an afterthought in the public-school curriculum (Al-Seghayer, 2019, p. 508). Furthermore, there has historically been a lack of theoretical, empirical, and practical research on English pronunciation issues in the Saudi classroom. While recent empirical studies have started to examine certain English pronunciation issues, they have not offered a comprehensive proposal to address those issues in the Saudi context (Alzinaidi & Abdel Latif, 2019). For instance, it is unclear which approach to teaching English pronunciation is more effective across language proficiency levels, which features should be prioritised to attain more intelligible speech based on realistic goals, when to start teaching pronunciation, and what the goals should be (e.g., attaining intelligible and comprehensible pronunciation).

Since the modern Kingdom of Saudi Arabia was founded in 1932, two major factors have led to the country promoting EFL in education. First, the oil industry has attracted foreign investors, most of them native speakers of English (Mahboob & Elyas, 2014). Second, the two holiest cities in Islam are in Saudi Arabia, and millions of Muslims make pilgrimage to the country each year to visit these cities. Since most of these visitors do not speak Arabic as a native language, English is often employed as a lingua franca (Alshahrani, 2016).

In addition to these historical and ongoing factors, in 2016, the Saudi government announced Saudi Vision 2030, a massive strategic vision for national development that seeks to reduce the country's dependence on the oil industry by reforming the public and private sector. A key element of this programme is improving education and preparing students for the 21st century job market. To do this, schools and universities are required to improve students' ability in English to better communicate with the rest of the world. Despite these factors

encouraging learning English, EFL education faces several challenges (Al-Seghayer, 2011). Any reform of the entire education sector will be difficult due to ongoing challenges to meet the goals of Vision 2030 (Alharbi, 2016).

Based on the lack of empirical research on teaching English pronunciation in the Saudi context, this chapter provides a review of English instruction in general. Furthermore, EFL instruction in Saudi Arabia has gained more interest from researchers, EFL teachers, experts, and policymakers in the past few decades (Alrashidi & Phan, 2015; Mahboob & Elyas, 2014). The following sections give an overview of teaching English in public school and higher education along with cultural factors. Subsequently, methods, techniques, and materials for teaching EFL are discussed. Finally, the chapter highlights immediate and essential reforms to teach EFL in the Saudi context.

## **2.2 The Saudi Education System and EFL Instruction**

In order to understand the challenges of EFL instruction in Saudi Arabia, it is important to examine the country's demographics, the structure of its education system, and the evolution of EFL instruction over time at different levels of education. Out of Saudi Arabia's population of 34 million, 11 million are expatriates (General Authority for Statistics, 2019). A considerable number of these expats use English to communicate with other residents. In addition, approximately 60% of the population is under 30. This relatively young population needs more education to pursue their careers as well as jobs when they finish their formal education. In education and business, English has taken on greater importance. Elyas and Picard (2010) pointed out that "English in particular is seen as a vehicle for economic development, as it is used as the *lingua franca* in large multi-national oil and gas companies" (p. 142). This is especially the case in Saudi Arabia, where the government has initiated a "Saudisation"

programme to increase the number of Saudis in the workforce, and the private sector is requiring communicative competence in English.

The Saudi education system can be divided into two formal systems: general and higher education (Ministry of Education, 2019). Under general education, there are three compulsory stages: elementary (six grades), intermediate (three grades), and secondary (three grades). There are also pre-school stages, such as kindergarten, which aim to prepare children for formal schools but are optional. Three types of schools in general education are available for students from age 6: public, private, and international. Public schools are free and strictly supervised by the government in terms of policies and curricula.

Before 2004, EFL was taught in public schools from the beginning of the intermediate stage (Al-Seghayer, 2011). In 2004, English became mandatory from sixth grade, and since 2011, students have started taking English in fourth grade. Although there have been serious attempts to teach English from first grade, there are several obstacles to make this goal more achievable. For instance, the instructional time allocated for English is still only two 45-minute classes per week at the elementary stage and four 45-minute classes per week in intermediate and secondary school (Al-Seghayer, 2014). Thus, the amount of time given for teaching English in public schools remains insufficient.

Private schools are similar to public schools in educational policies, standards, and curricula (Al-Seghayer, 2011). For instance, English is not the medium of instruction in the majority of private schools. Furthermore, they are similar to public schools in terms of time allotted for teaching English and the textbooks used.

International schools are completely different (Al-Seghayer, 2011). English is the formal medium of instruction in most international schools for two reasons. First, students belong to different L1 backgrounds, not only Arabic. Consequently, they need to use English

to communicate with each other. Second, international schools have adopted American, British, Canadian, or Australian curricula, which requires using English as a medium of communication and instruction to increase education quality (Alrashidi & Phan, 2015; Elyas & Al-Grigri, 2014). Although international schools report better educational outcomes than public and private schools, it is unclear whether using English as the formal medium of instruction is related to this success.

Saudi higher education includes three types of institutions. The first is the university. There are 29 public and 14 private universities that offer a range of scientific, social, and business programmes for bachelor's, master's, and doctoral degrees. Although public universities are fully subsidised by the government, they are partially independent in terms of administrative and academic affairs (Ministry of Education, 2019). The number of universities has increased in the last decade because of the growing demand of the Saudi labour market and the rising number of students ready to enter a university (Alharbi, 2016). The second type is the community college, which offers a two- or three-year programme in certain majors (Ministry of Education, 2019). Such colleges are affiliated with one of the local public universities and offer alternative options for students who cannot join a university. The third type is the technical and vocational college, which offers specific programmes associated with labour market needs.

To help students transition from secondary to higher education, students go through a Preparatory Year Program (PYP) in public and some private universities (Alasmari, 2013; Alghamdi, 2015).<sup>4</sup> In addition, English in the past was not offered as a required subject or the formal medium of instruction in the majority of public and private universities. Now, however, university students are required to study English at least as a mandatory subject in the PYP, including students who are going to study their majors in Arabic as a medium of instruction.

---

<sup>4</sup> Although private universities are under the supervision of the Ministry of Education, some have their own policies and regulations.

The PYP principally aims at preparing new undergraduate students to be more academically and psychologically ready to cope with the university environment and its expectations.

There are three main streams in the PYP, each offering specific intensive subjects during two regular semesters (Ministry of Education, 2019). Some subjects might be offered for exceptional cases in a summer semester. The Health stream seeks to prepare students for medicine, applied medical sciences, pharmacy, and dentistry colleges. English classes in this stream are intensive, taking up 20 hours per week. The Engineering and Sciences stream prepares students for applied and theoretical science colleges, including computer information sciences and a wide range of engineering tracks. The allotted time for teaching English in this stream is 20 hours per week. The Humanities stream prepares students for educational and humanities colleges, including arts, education, law and political science, languages, and translation studies. The amount of time devoted to English in this stream is 16 hours per week. This is less than other streams because some majors in the humanities are taught in Arabic (Alasmari, 2013; Alghamdi, 2015).

Although the PYP has attempted to address the unsatisfying learning outcomes of general education, two critiques have been raised (Alamri, 2011). First, the distribution of streams needs to take into account that each major has its own status. Combining majors under the same stream might make the PYP less helpful for students in some majors. A further observation is that no explicit theoretical and pedagogical guidelines have been identified as a standard for assigning the curriculum of required subjects, in particular English.

To better understand the state of English instruction in a non-native context, cultural factors should be considered (Elyas & Picard, 2010). This is because a complex set of societal and cultural factors have affected the implementation of English instruction in the Saudi education system. In the past, for example, some families discouraged their children from

learning English, fearing it would corrupt their values (Alrashidi & Phan, 2015; Ullah, 2017). Now, however, even more conservative families encourage their children to learn English, which they see as a gateway to a better future by equipping students to meet the shifting needs of the Saudi job market (Alrashidi & Phan, 2015).

This change has led policymakers to pay more attention to developing the English curriculum and increasing instructional time in public education. However, cultural concepts in mainstream English curriculum represent a challenge because many Saudis fear any curriculum adopted from English-speaking cultures would lead to more Westernisation of Saudi society. As a result, “Despite encouragement from government and industry to expand the teaching of English, a hostile attitude has persisted towards English in some sectors of Saudi Arabian society” (Elyas & Picard, 2018, p. 71). However, teaching English pronunciation to achieve intelligible and comprehensible speech among the Saudi EFL learners has not been a central issue for the policymakers throughout the school and tertiary levels (Al-Seghayer, 2019).

### **2.3 Methods, Techniques, and Materials for Teaching EFL**

When English first started being taught in Saudi Arabia, the methods, techniques, and materials were adopted from American and British contexts (Al-Hajailan, 2006; Al-Seghayer, 2011). The evolution of teaching methodology can be divided into three stages (Al-Hajailan, 2006). In the first stage, the grammar-translation method was adopted to improve reading and writing by having students translate English content into Arabic. The only technique employed in this stage was to ask students to memorise the target vocabulary through repetition.

Due to criticism of this approach and major changes in the Saudi education system in the mid 1950s, the audio-lingual method was implemented to keep pace with theoretical and methodological trends at the time (Al-Hajailan, 2006). The primary goal of this method was to develop oral and listening skills through drills and repetition and explain the basic English

grammar structures presented in short dialogues. Although this method yielded better outcomes, it had its own problems (Al-Hajailan, 2006). For instance, it was not appropriate for a non-English-speaking environment due to the lack of opportunities to practice the short dialogues outside the classroom.

As an emerging approach in the field of SLA/FLA during the 1970s and 1980s, the communicative language teaching (CLT) method was adopted to develop communicative competence in the Saudi EFL classroom (Celce-Murcia et al., 1997). In other words, students acquire L2 communicative skills based on real-life situations. Various techniques could engage students in real communicative conditions, such as problem-solving and roleplaying activities.

In the Saudi context, the main goal of using this approach was to shift the emphasis of teaching from theoretical and pedagogical grounds to using English as a communication tool in daily life. Consequently, Saudi EFL students were expected to learn how to effectively communicate and interact in English (Al-Hajailan, 2006). Although this method appeared more practical by integrating the four basic skills (speaking, listening, reading, and writing), the instructional time allocated for EFL has remained insufficient to achieve the goals of this approach.

Another major constraint is a lack of training for EFL teachers on how to implement the theoretical and practical concepts of CLT in the classroom (Al-Seghayer, 2011). Before the government increased the number of higher education institutions in Saudi Arabia, the same methods and techniques for teaching English were employed at all levels without considering their appropriateness for each level of education.

Throughout the last six decades, the EFL curriculum in Saudi Arabia has been frequently modified to solve ongoing difficulties in the education system or to meet new challenges (Al-Hajailan, 2006). Originally, no textbooks were specially designed for Saudi EFL

classes. Two books were used to develop reading and writing skills, *Red Primer* and *Green Primer*, for the primary and secondary schools,<sup>5</sup> while *Readers 1, 2, 3, 4, and 5* supplemented the secondary stage (Al-Hajailan, 2006).

In the late 1950s and early 1960s, Saudi curriculum, including English, witnessed critical and comprehensive reform (Al-Seghayer, 2011). In cooperation with the Longman publishing company, a new textbook, *New Living English for the Arab World*, was formally presented for the intermediate stage. Although this book was used for over two decades, it did not meet Saudi EFL learners' interests and objectives and failed to integrate English skills with Saudi socio-cultural norms (Al-Hajailan, 2006). The textbooks *Living English Structure* and *First Book in Comprehension and Practice* were used for the first year of the secondary stage, and *Vice Versa* was used in the second year. *Easier English Practice* and *All Around the World in Eighty Days* were used for students in the third year.

To address the failure of these textbooks to increase learners' communicative competence, a new textbook was launched in cooperation with Macmillan Press, *Saudi Arabian Schools English* (Al-Seghayer, 2011). Different editions for each year in the intermediate and secondary stages were supplied. Although this textbook did not display crucial educational improvements among Saudi EFL learners, it was used from the beginning of the 1980s to 1995. Its unsuccessful implementation shows a failure to link pedagogical foundations to learner goals and needs (Al-Seghayer, 2011).

In an attempt to localise the design of English textbooks, the Saudi Ministry of Education in cooperation with English teaching experts in the Saudi context introduced a new series of textbooks in 1995, *English for Saudi Arabia* (Al-Seghayer, 2011). One of their advantages was an emphasis on Saudi socio-cultural values by communicating in English about

---

<sup>5</sup> There were only primary and secondary schools at that time, and secondary school was a five-year period. Secondary school was then divided into two stages, intermediate and secondary, each lasting three years.

real situations that students encounter every day. However, this textbook was not implemented with a computer-based approach.

In 2004, English became mandatory from the sixth grade, and a new series of textbooks was launched, *Say It in English*, which aimed to improve Saudi EFL learners' communicative competence in relation to new local and global challenges (Al-Seghayer, 2011). Although this textbook met the expected objectives, further national projects were initiated in 2008, 2011, and 2013 to develop the Saudi curriculum, including English materials (Alyami, 2014; Mitchell & Alfuraih, 2017). For example, Tatweer is a national company responsible for developing curriculum with the Ministry of Education. These developments have aimed for more improvement in English communicative skills (Alrasheed, 2010).

Methods, techniques, and materials for teaching English in the Saudi context have undergone a number of developments. However, there is still much debate about how to teach English effectively given that classroom outcomes remain unsatisfying (Mitchell & Alfuraih, 2017). This section focused on public schools since the same issues have applied to private schools and universities. Regarding curriculum, private schools and universities use a different variety of popular international textbooks, such as *Headway*, *Cutting Edge*, and *Face2Face* (Alghamdi, 2015; Al-Seghayer, 2011). In contrast, international schools are fully independent, and each school uses different methods, techniques, and materials based on the objectives of the education system of the institution (e.g., American or British) (Al-Seghayer, 2011). However, public and private schools and governments do not address pronunciation issues and do not offer an explicit approach to teaching English pronunciation in order to increase intelligibility and comprehensibility in the Saudi EFL classroom (Al-Seghayer, 2019). In addition, Saudi English course materials have not explicitly included pronunciation aspects. Thus, it is difficult to review pronunciation in these materials (Al-Seghayer, 2019, p. 508).

## 2.4 Essential Reforms for Teaching EFL in the Saudi Context

Since teaching EFL has been formally implemented in Saudi Arabia, many experts, EFL teachers, and policymakers have called for reforming the status of EFL instruction (Alqahtani, 2018). As a necessary step to reform, the Ministry of Education launched a new teaching framework in 2013 that seeks to develop students' communicative competence in English by focusing on a number of objectives:

1. Explain the tenets of Islam with a vision to promoting international understanding and tolerance.
2. Advocate and participate in spreading Islam.
3. Promote mutual cultural understanding and respect among nations.
4. Enhance their cognitive and problem-solving skills, thus leading to academic and professional advancement.
5. Develop an awareness of the significance of English as a means of international communication.
6. Develop a positive attitude towards learning the English language.

(Ministry of Education, 2013, p. 9)

An obstacle to these objectives is that insufficient time has been allocated for English instruction, especially in public education, for two reasons (Al-Seghayer, 2011, 2014). First, students are not exposed to English outside the classroom, and four 45-minute classes a week for intermediate and secondary school and only two a week for primary school is not enough to teach pronunciation and make it a part of English class due to time constraints. Hence, allocating more time for English is necessary to improve communicative competence.

Another problem is student motivation. As stated in a number of empirical studies (e.g., Al-Seghayer, 2011; Mahdi, 2015; Rahman & Alhaisoni, 2013), students lack intrinsic motivation to learn English as they believe English is not important in their daily lives outside the classroom. This is because they live, study, and often work in a monolingual environment where they only need Arabic. This lack of motivation cannot be remedied by simply increasing the time dedicated to English or changing the materials. Students should be motivated to learn English by triggering their inner goals. In addition, Saudi EFL students have experienced

English as an academic subject they are simply asked to pass. Almari (2011) pointed out that “higher education in Saudi Arabia is missing important elements, which depend on either hygiene or motivation factors in motivating the students through the learning process to achieve the objectives of the programs offered” (p. 90). Therefore, the importance of learning English must be connected to students’ goals outside the classroom.

Most if not all English teachers in Saudi public and private schools are non-native English speakers and require pre- or in-service training. Although intensive English programmes in Saudi universities recruit native English-speaking teachers, these teachers also need training because a native-speaking performance is not sufficient to be qualified to teach EFL (Ahmad et al., 2016). Thus, both native and non-native teachers need pre- and in-service training to be effective (Al-Harbi, 2006; Al-Hazmi, 2003; Althobaiti, 2017). For instance, teachers should be trained on how to use the communicative competence method to develop students’ English skills and how to use new methodologies, materials, and technology. Additionally, they should be aware of students’ different learning styles and individual differences and should possess adequate pedagogical and theoretical knowledge.

A number of studies (e.g., Abu-Ghararah, 2014; Aljumah, 2011; Al-Seghayer, 2014; Hamad, 2013; Khan, 2013) have argued for a range of strategies and activities, such as repetition-based drills, roleplay, group discussion, and jigsaw puzzles. However, pronunciation has not been a central point of research in the Saudi EFL context (Al-Seghayer, 2019). Additionally, the new framework has given no explicit proposal regarding the extent to which Saudi EFL learners’ intelligibility and comprehensibility should be improved. The importance of English segmental and suprasegmental features in teaching has not been clarified or which features might enhance intelligibility and comprehensibility. Last but not least, it is unclear

whether a certain type of instruction (e.g., explicit instruction) could make Saudi EFL learners more intelligible and comprehensible.

## **2.5 Chapter Summary**

This chapter gave an overview of EFL instruction in Saudi Arabia. The best methods of instruction have been debated for decades, but many issues have been underexplored. As a result, much uncertainty exists about the effectiveness of methods and materials for teaching English, leading to ongoing calls for reform. The next chapter reviews the literature on a variety of issues relevant to this study, including L2 pronunciation instruction.

## **Chapter 3: Literature Review**

### **3.1 Introduction**

This chapter provides a critical review of the L2 pronunciation literature, starting with the theoretical framework, segmental and suprasegmental aspects of English pronunciation, and the concepts of intelligibility and comprehensibility. It goes on to cover a range of issues in L2 pronunciation instruction, such as intelligibility and nativeness in pronunciation teaching, the effect of instruction on segmentals and suprasegmentals regarding intelligibility and comprehensibility, the effect over time, and the effect on beginning-level learners. After that, it reviews the research on the relationship between motivation and the attainment of more intelligible and comprehensible speech. It then turns to empirical studies that have explored L2 learners' perceptions of the effectiveness of pronunciation instruction and studies on how native speakers of English perceive non-native pronunciation. Based on the previous literature, the chapter states the research questions.

### **3.2 Theoretical Framework of L2 Phonology**

Numerous L2 speech models have been formulated to describe, explain, and predict the nature of L2 sound acquisition. Nevertheless, questions remain, such as the interference between L1 and L2 sound systems, which L2 developmental processes take place over the course of speech acquisition, which difficulties are mastered, and what sounds are modified.

Several proposed theories have failed to explain a set of complex issues in L2 phonology. For instance, the Contrastive Analysis Hypothesis was one of the first theories to attempt to explain the nature of L2 sound acquisition but was ultimately rejected because it focused solely on using the characteristics of the L1 sound system to produce L2 sounds (Flege, 1995; Best, 1995; Derwing & Munro, 2015). Furthermore, some L2 speech theories lack pedagogically useful information (Derwing & Munro, 2015, p. 63). This section provides a critical review of

the most influential L2 speech models, namely the Speech Learning Model (SLM), Perceptual Assimilation Model (PAM), and PAM-L2.

These models were selected for review because they could provide explicit theoretical and practical descriptions, predictions, and explanations of the cross-linguistic influence of mother language sound characteristics on target language sound production accuracy (Colantoni et al., 2015). In addition, this was a classroom-based study, and the SLM and PAM could help inform how to teach pronunciation in class more effectively. Derwing and Munro (2015) supported this as follows:

Approaches like PAM and SLM that identify acquisition mechanisms have the potential to inform teaching practices. For instance, the link between perception and production posited in SLM strongly implies that perceptual skills should be taught as part of pronunciation instruction. (p. 69)

### **3.2.1 SLM**

SLM claims learners perceive L2 sounds based on their native phonological system. L2 sounds might be accurately produced by incorporating L2 articulatory patterns into L1 phonetic categories. Learners use a mechanism called equivalence classification to classify L2 sounds as either equivalent or not equivalent to the L1 sound categories. Therefore, learners incorporate L2 sounds similar to their L1 into L1 phonetic categories and establish a new phonetic category for L2 sounds perceived differently from the L1 phonetic categories (Flege, 1995; Flege & Bohn, 2021; Flege et al., 2003; Guion et al., 2000).

SLM was constructed to address L2 production issues among adult L2 learners but “claims that without accurate perceptual targets to guide the sensorimotor learning of L2 sounds, production of the L2 sounds will be inaccurate” (Flege, 1995, p. 238). SLM’s theoretical foundations, as presented by Flege (1987), claim that new L2 sounds dissimilar to L1 sounds should be easier to acquire, while L2 sounds similar to L1 sounds should be more difficult.

Building on this assumption, SLM offers four major theoretical proposals (Flege, 1995). The first assumes the processes and categories of the L1 sound system could be used in learning the L2 sound system. The second is that phonetic categories represent certain features of speech sounds in each language. This means each language has its own sound characteristics. The third states that phonetic categories formed for L1 sounds would be gradually developed to determine the L1 and L2 sound properties. The fourth postulates that differentiating L1 and L2 phonetic categories requires enormous effort from bilingual learners, because these categories share the same phonological space in the mind.

SLM bases predictions on seven hypotheses (Flege, 1995). The first (H1) states that the perceptual relation between L1 and L2 sounds is not only on an abstract phonemic level but on an allophonic level as well. The second (H2) claims that learning L2 sounds could, as a result of realising phonetic differences, establish a new category to phonetically distinguish between L2 sounds and the nearest L1 sounds. The third (H3) claims it is likely easier to perceive dissimilar L2 sounds as the nearest L1 sounds because the phonetic differences between L1 and L2 sound systems are obvious. The fourth (H4) posits that as the age of learning increases, it becomes more difficult to discern phonetic differences between L1 and L2 sounds, especially L2 sounds that are not contrastive with L1 sounds. Numerous studies have emphasised the importance of age of learning in foreign or native accent formation (e.g., Flege & MacKay, 2004; Piske et al., 2001).

The fifth hypothesis (H5) states that equivalence classification is probably connected to L1 and L2 diaphones that would not allow the creation of a separate L2 perception system (Flege, 1995). Therefore, one perception system might be used for L1 and L2 sounds, which could influence the production of L1 and L2 sounds separately. For example, Flege and Port (1981) investigated whether Saudi speakers of English could produce a set of minimal pairs of

English plosives, targeting those absent from Arabic, such as /p/, compared to those shared between the two, such as /t/ and /k/. The acoustic measurements were equally authentic, whereas a perception task revealed native English speakers discerned Saudi speakers' /p/ to be as accurate as the English phonemes /t/ and /k/.

The sixth hypothesis (H6) suggests bilinguals are able to create a different phonetic category from their L1s, but this could be different from monolingual categories because of the influence of bilingual L1 categories (Flege, 1995). For instance, Munro (1993) found that Arabic-English bilinguals who had lived in the U.S. for almost six years made an enormous effort to produce the English vowels /i/ and /ɪ/ like native speakers, as demonstrated by spectral and temporal exaggeration. The seventh hypothesis (H7) claims L2 learners eventually produce L2 sounds in correspondence with the properties in their phonetic categories (for further details, see Flege, 1995, pp. 239–243).

Although SLM helps researchers obtain a more profound theoretical understanding of L2 speech acquisition, it has shortcomings. For example, it focuses solely on acquiring English phonetic sounds in English-speaking countries, such as Canada. If SLM hypotheses were examined for learners in non-English-speaking countries and classrooms, the acquisition of phonetic sounds could offer new directions in L2 phonology. Flege and Bohn (2021) stated that “The formation of new phonetic categories for L2 sounds will depend on the quantity and quality of L2 input” (p. 3). However, it is unclear what type of input this would be, whether it would be received in a naturalistic or structured environment, and whether the ultimate goal of input is to achieve intelligible and comprehensible speech or native-like output.

In the present study, SLM has been used to address three theoretical aims. The first was to explore whether age of learning is a crucial factor helping Saudi late learners achieve intelligible and comprehensible speech in English. Such learners might not realise the phonetic

differences between English and Arabic sounds, particularly the non-contrastive ones, as argued in PAM H4.

The second was to explore whether Saudi beginner learners could establish a separate English phonetic category after being exposed to pronunciation instruction, which would increase their perceptual awareness of specific phonetic differences between Arabic and English sounds. Based on PAM H6, it is unclear whether Saudi learners will create a new English sound category identical to an English monolingual category because of the influence of the Arabic category.

Third, equivalence classification is essential because Arabic and English sounds might be phonetically and acoustically associated, which could lead to Saudi EFL learners using only one phonetic category for two different sound systems. The study sought to explain whether this occurs due to phonetic and acoustic similarities between Arabic and English sounds, as postulated in PAM H5.

### **3.2.2 PAM and PAM-L2**

PAM investigates the differences between non-native contrasts in terms of cross-language perception. It focuses on how non-native listeners perceive L2 sounds on the basis of their (dis)similarities to an L1 (Best, 1995; Best & Tyler, 2007). PAM proposes a theoretical framework for how non-native contrasts are perceived under the influence of a native perception, which plays a central role in perceiving non-native phonemic contrasts (Best, 1993, 1994a, 1994b, 1995; Best et al., 1988).

PAM was built on the ecological theory of perception, also called the direct-realist approach or articulatory phonology (Browman & Goldstein, 1989, 1992; Fowler, 1986, 1989; E. Gibson, 1991; J. Gibson, 1979). These approaches argue that the development of native speech perception occurs as a result of infants' perceptual abilities forming based on speech

signals in the vocal tract, which have a substantial effect on realising the phonetic details of non-native contrasts.

According to PAM, non-native listeners perceptually categorise a new L2 phoneme as the most similar L1 phoneme (Best, 1993, 1994a, 1995; Best et al., 2001). If non-native listeners categorise an L1 phone well, it is considered to be categorised; when this does not occur, it is called uncategorised. When the sound has not been assimilated as a sound, it is called non-assimilated.

PAM provides a set of predictions for how non-native listeners categorise L2 sounds as an example of categorisation of a single non-native phone to a native phoneme (Best & Strange, 1992). In particular, if non-native listeners optimally categorise L2 phones, it is called two-category (TC) assimilation. For instance, Japanese speakers of English assimilate English phones /w/ and /j/ separately, which is considered an optimal example of TC assimilation because these sounds correspond to Japanese /w/ and /j/. On the other hand, if L2 sound are well or poorly categorised, it is called single-category (SC) assimilation. For instance, Saudi speakers of English, particularly at the beginning level, poorly categorise English /n/-/ŋ/ contrasts because the phoneme /ŋ/ does not exist in the Arabic phonemic inventory; therefore, such a categorisation would be considered SC (Alharbi, 2014).

The last scenario of L2 phone categorisation, when non-native listeners intermediately categorise L2 phones, is called category-goodness (CG) assimilation (Best, 1994a, 1995). PAM's primary prediction is that "the discrimination performance pattern for adults should be, from highest performance to lowest: TC > (NA<=>CG) > SC" (Best, 1994a, p. 182).

Regarding categorisation of two or more non-native phones to native phonemes, when one of the L2 phones is categorised to one native phoneme while another phone is not, this is called uncategorised-categorised (UC) assimilation (Best, 1994a, 1995). Conversely, when both

non-native phones are not categorised into native phonemes, this is called uncategorised-uncategorised (UU) assimilation. It is rare for L2 phones not to be assimilated into native phonemes due to the different bases of articulations of L1 and L2 phonemes. This case is called non-assimilable (NA).

PAM-L2 examines the most influential models (PAM and SLM) in L2 phonology by comparing the ability of L2 learners to perceive L2 contrasts and non-native contrasts (Best & Tyler, 2007). To conduct a theoretical comparison, PAM-L2 makes its predictions by reviewing four main SLM predictions from a PAM standpoint. First, L2 learning is linguistically formed based on the L1 learning process. PAM-L2 shows a partial agreement with this prediction because PAM-L2 suggests that L2 perceivers realise articulatory gestures relying on speech signals, not acoustic-phonetic cues as SLM claims. Second, speech sounds have language-specific features that form phonetic categories in long-term memory. PAM-L2 does not support the claim that mental representations play a role in L2 perceptual learning, because PAM-L2 assumes L2 learners can use different perceptual devices to transform L1 sound characteristics into an L2. Third, phonetic categories in the early stages of L1 sound acquisition gradually develop to determine L1 and L2 phone features. PAM-L2 agrees with this prediction, positing that listeners can develop and modify their perceptual abilities over time. Fourth, due to the existence of “L1 and L2 phonetic categories in a common phonological space” (Best & Tyler, 2007, p. 22), bilinguals might encounter difficulty always differentiating between L1 and L2 phones. PAM-L2 offers partial support for this prediction, pointing out that L2 speech learning occurs due to the interaction of phonetic and phonological levels, whereas SLM merely discusses the phonetic level (see Best & Tyler, 2007, pp. 15–24). Figure 3.1 illustrates the outline of PAM/-L2 patterns and their predictions about similarities between L2 and L1 sounds.

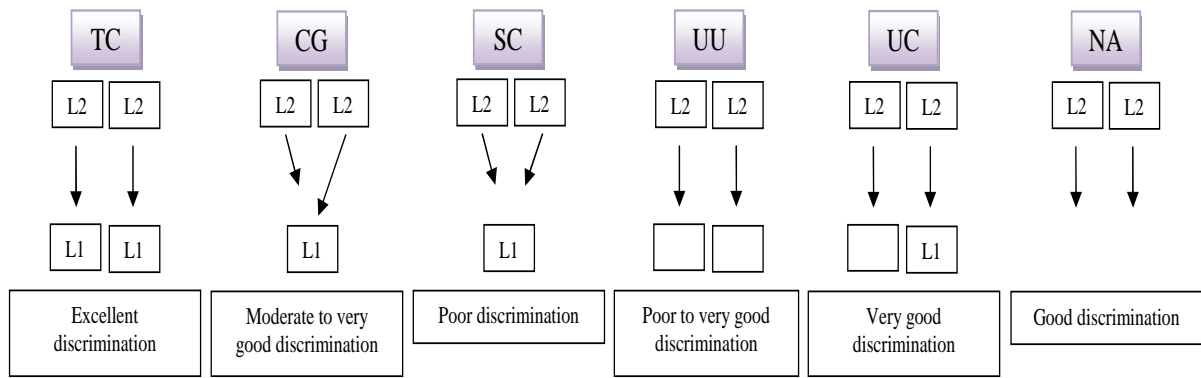


Figure 3.1. Outline of PAM/-L2 patterns and predictions based on L1-L2 similarities (adopted from Almbark, 2012, p. 47).

Best and Strange (1992) suggest the amount of L2 exposure can reorganise the patterns of L2 perceptual assimilation. However, it is unclear what L2 perception mechanisms are working and their influence on L1 perception (Escudero, 2007). Thus, PAM/-L2 fails to argue whether the same phonemic developmental changes in speech perception among adult non-native learners would occur when perceiving non-native phonemic contrasts. It is also unclear whether phonemic developmental changes would occur because of native phonemic patterns influencing the perceptual success of non-native phonemic contrasts.

Drawing on Best and Tyler (2007), non-native perceptual developments, among young or adult learners, are beyond the scope of PAM/-L2, which broadly concentrates on cross-linguistic speech perception in natural non-native contexts, not in a foreign classroom. Thus, a gap that has not been addressed from a PAM/-L2 approach is how Saudi EFL learners assimilate English sounds (with poor or good discrimination) after being exposed to pronunciation instruction.

While some PAM/-L2 hypotheses have been investigated in naturalistic settings, the current study was conducted in a foreign-language classroom. However, PAM/-L2 is still helpful in this context. For instance, it might explain which type of pronunciation instruction could help learners improve their perceptual discrimination of English contrasts as a TC pattern.

### **3.2.3 Comparisons between SLM and PAM/-L2**

SLM explores how L2 learners, especially adults, perceive L2 sounds in a native-like fashion (Flege & Bohn, 2021). In contrast, PAM examines the perception of non-native contrasts among native and non-native listeners. PAM-L2 is an extension of PAM used to address L2 acquisition and compare L1 and L2 sound categorisations. Although PAM-L2 agrees with the SLM principles that L1 and L2 categories share the same phonological space, it is different from SLM as the accuracy of L2 perception depends on a degree of perceptual assimilation, such as TC, SC, or CG assimilation, while SLM depends on equivalence classification.

In early L2 learning, SLM and PAM/-L2 suggest that L1 sounds could be categorised as L2 sounds because the L1 and L2 are integrated systems (Best & Tyler, 2007; Flege & Bohn, 2021). In the next stage, although SLM claims L2 learners could access L2 development and create L2 categories, there is no formal proposal to explain whether the same L1 mechanisms would be used. Similarly, PAM/-L2 postulates that L2 categories would be reorganised from L1 categories, but it is unclear which mechanisms would be employed. Finally, SLM considers the ultimate attainment of L2 perception to be complex because it depends on age of learning and the integration of L1 and L2 perception in the same phonological space. PAM/-L2 has no formal proposal for the end state of L2 speech perception since it does not explain the mechanisms in this state or predict the possible outputs.

Despite their influence in L2 phonology, SLM and PAM are segment-based theories for the acquisition of L2 speech. Thus, “the two models remain narrow in their scope, targeting only the learning of individual sounds” (Foote & Trofimovich, 2018, p. 78). Due to the lack of a comprehensive theory, Colantoni et al. (2015) asserted that “None of the theories of L2

phonology to date make explicit predictions about the acquisition of prosody” (p. 312). Hence, there is a need for a theoretical framework to describe and predict this acquisition.

### **3.3 Segmental and Suprasegmental Aspects of English Pronunciation**

#### **3.3.1 English Segmentals**

This section gives an overview of English segmentals, as they are crucial for intelligible, comprehensible speech (Derwing & Munro, 2015; Levis, 2018). In L2 perception, the importance of segmentals stems from their role in making L2 speakers more understood once they are sufficiently accurate in producing individual sounds. In other words, L2 speakers can be better understood when they produce L2 sounds in a way that listeners expect to hear. However, sufficient accuracy does not mean a perfect match, which is an unrealistic goal (Levis, 2018).

Levis (2018) made a number of important arguments about segmentals in L2 pronunciation (p. 62). First, segmentals are highly important for intelligibility and comprehensibility as they determine whether the correct word is perceived. Second, segmental errors are not equal in terms of causing misunderstandings. Third, segmentals are influenced by different types of mispronunciations, and they have varied effects on understanding L2 speakers. The following subsections describe how English sounds are classified. Then, common English pronunciation errors are classified, particularly English pronunciation errors frequently made by Saudi EFL learners.

##### **3.3.1.1 Classification of English Segmentals**

Given this study’s focus on consonants, only English consonants are described. The English sound system consists of 24 common consonants (Rogerson-Revell, 2018). Table 3.1 presents the English consonant phoneme inventory from Rogerson-Revell (2018, p. 100). The consonants described here are based on General British and North America English. However,

there are differences in how sounds are produced due to allophonic variation, the position of phonemes in a syllable, regional variation, and other factors.

Table 3.1: English Phonemic Consonant Inventory

	Bilabial		Labio-dental		Dental		Alveolar		Palato-alveolar		Palatal	Velar		Glottal
Plosive	p	b					t	d				k	g	
Fricative			f	v	θ	ð	s	z	ʃ	ʒ				h
Affricate									tʃ	dʒ				
Nasal		m						n					ŋ	
Approximant		w							ɹ		j		(w)	
Lateral Approx.								l						

English has six plosives in three voiced/voiceless pairs, although aspiration can differentiate between these pairs (Rogerson-Revell, 2018). For instance, the voiceless plosives are aspirated word-initially, while voiced plosives are not. The difference in aspiration appears to be more significant than the absence or lack of voicing. As a result, L2 learners (e.g., L1 speakers of Arabic) should be more aware of the importance of aspiration in English, especially if their L1 does not allow initial aspirated plosives. English has five pairs of voiced/voiceless fricatives and one glottal fricative /h/. The two affricates, /tʃ/ and /dʒ/, can pose a challenge as English learners sometimes replace them with a fricative. For instance, Saudi EFL learners might replace /tʃ/ with /ʃ/ to pronounce “much” (Rogerson-Revell, 2018). Of the three nasals, /m, n, ŋ/, /ŋ/ is more problematic for EFL learners as it is less common in other languages, including Arabic. Regarding approximants, /ɹ/ is challenging because it is pronounced differently across dialects. According to typical descriptions of British English, /ɹ/ is a post-alveolar approximant only word-initially (Rogerson-Revell, 2018, p. 104). Certain sound patterns also affect English consonant production, such as glottalisation, flapping, and rhotic versus non-rhotic accents.

### 3.3.1.2 Common English Segmental Errors among Saudi EFL Learners

Several English consonants have no equivalent in many varieties Arabic and thus represent a problem for speakers of those varieties (Binturki, 2008). Due to this variation, English pronunciation errors can differ depending on the Arabic dialect learners speak and the society they come from (Barros, 2003). Table 3.2 shows the English consonants with no equivalent in the Northern dialect of Saudi Arabic. While /g/ exists in some Arabic dialects, the current study was carried out in Aljouf University and all participants were native speakers of the Northern Saudi dialect.

Table 3.2: English Consonants without a Northern Saudi Arabic Equivalent and Common Substitutions

English Consonant	Common Substitution
/p/	/b/
/g/	/dʒ/ or /ʒ/
/v/	/f/
/tʃ/	/ʃ/
/ŋ/	/ng/, /nk/, or /n/

### 3.3.2 English Suprasegmentals

This section describes the features of English prosody. Suprasegmental features provide clues to determine the important parts of a speaker's message (Derwing et al., 1998; Field, 2005; Hahn, 2004; Zielinski, 2008), thus playing a fundamental role in achieving more intelligible and comprehensible pronunciation (Levis, 2018). Furthermore, the intelligibility of segmentals often relies on suprasegmentals, such as stress placement in individual words and the rhythmic patterns in sentences and phrases.

Highlighting the importance of prosodic features in L2 speech perception and production, Murphy (2017) pointed out that "prosody represents the phonological context in which all other facets of L2 pronunciation occur" (p. 32). As a result, Setter and Sebina (2018)

called for more attention on learning prosody to achieve intelligible and comprehensible pronunciation.

Some major suprasegmental features include word stress, rhythm, thought group, connected speech (e.g., linking), focus (also called prominence), and final intonation. In L2 learning, the importance of suprasegmental features depends on learners' L1 background, language proficiency, and learning context (Murphy, 2017). The following subsections describe English lexical stress, rhythm, and intonation as they have been found to be more challenging for Saudi EFL learners (Swan & Smith, 2001).

### **3.3.2.1 Lexical Stress**

In theoretical terms, lexical stress is when greater force of articulation is placed on certain syllables in a word to provide useful linguistic information (Setter & Sebina, 2018). Intensity or loudness help determine where the stress is placed. There are generally two types of lexical stress: primary and secondary. Lexical stress in English occurs in words that have two or more syllables. The primary stress falls on one syllable, which is longer in duration, is higher in pitch, and has a full vowel (Cutler, 2015). More importantly, lexical stress placement changes the category and meaning of words. For instance, “project” is a noun when stress is on the first syllable (“PROject”) but a verb when stress is on the second syllable (“proJECT”). Thus, “lexical stress does provide highly useful cues to listeners as to where word boundaries are to be located in speech signals” (Cutler, 2015, p. 120). Misplaced stress in English can reduce intelligibility and comprehensibility for native and non-native speakers (Field, 2005). This is because “When a word, especially a word central to the understanding of the message, cannot be recognised, listeners may stop all other processing to decode the word that was not understood” (Levis, 2018, p. 100).

Arabic has three types of lexical stress: primary, secondary, and weak (Swan & Smith, 2001). Like English, stressed syllables in Arabic are longer in duration and higher in pitch. However, the rules in Arabic that determine which syllable is stressed are different from English. For instance, only one of the last three syllables in an Arabic word can be stressed, but this depends on the syllable and whether the word has long and short vowels (Swan & Smith, 2001).

Although Arabic and English are both stress-timed languages, English has free stress, meaning stress can be placed on any syllable of a word with more than one syllable, while Arabic is a fixed-stress language, in which stress placement is regular and predictable (Swan & Smith, 2001). As a result, word stress in English can be problematic for Saudi EFL learners. However, it is unclear whether different dialects of Arabic have different lexical stress patterns that affect the accuracy of EFL lexical stress in different ways.

Saudi EFL learners encounter substantial difficulties placing the correct stress in English words, which could impede communication when listeners fail to recognise a word (Levis, 2018). Nevertheless, not all lexical stress errors are equal in their impact on intelligibility and comprehensibility.

### **3.3.2.2 Rhythm**

The timing patterns of a language can be defined by rhythm in the form of alternating strong and weak beats (Low, 2014). Early research attempted to identify the characteristics of rhythm by determining which speech unit consistently recurs, such as division of time into equal parts. For instance, Dauer (1983) assumed the different rhythmic patterns between stress-timed and syllable-timed languages could be identified based on differences in syllable structure, vowel reduction, and stress rules. However, recent research has redefined rhythmic patterns (Low, 2015).

Low (2014) stated that rhythm is largely influenced by syllable length, which is affected by a number of factors, such as stress, which could give more prominence to a particular syllable by changing length, loudness, pitch, and vowel quality. Another factor is accent, as prominence is induced mainly by pitch obtrusion that often occurs on a syllable in an utterance that carries important information. A third factor is when the duration of full vowels is longer than reduced vowels. A fourth is syllable structure as the durations between open and closed syllables differ, as longer durations can be found in open syllables (ending with vowels) than closed ones (ending with consonants). A fifth factor is the placement of pauses, which influence rhythm when they occur within or after an utterance.

As noted by Celce-Murcia et al. (2010), stress rules in English rhythm are based on word categorisation. For instance, content words (i.e., nouns, verbs, adjectives, adverbs, wh-words, negatives) are often stressed because they carry more important information, while function words (i.e., articles, prepositions, pronouns, auxiliary verbs) are often unstressed or weakly stressed.

Based on the above conceptual and theoretical overview of rhythmic patterns, the ultimate goal of rhythm is to help listeners divide continuous speech to determine where important words begin and end. Levis (2018) indicated that rhythm might affect the understanding of speech by listeners (intelligibility) and the degree of ease when listeners attempt to process speech (comprehensibility). Furthermore, “rhythm and related features, especially speech rate (a correlate of spoken fluency), appear to affect judgments of accentedness and comprehensibility, as well as impacting intelligibility” (p. 143). Saudi EFL learners encounter challenges with English rhythmic patterns because unstressed syllables tend to be more reduced in English, while Arabic usually shows no major differences in the force of

stressed and unstressed syllables. This means even an unstressed syllable in Arabic can have a full vowel (Kenworthy, 1987).

### **3.3.2.3 Final Intonation**

Although prosodic features have received growing attention, more research is needed as “Little research exists about how intonation is acquired by adult learners or the extent to which cognitive constraints play a role during the acquisition process” (Wennerstrom, 2018, p. 154). Addressing the need for theoretical and empirical descriptions of L2 intonation learning, Mennen (2015) formulated the L2 intonation learning theory to propose a working model of L2 prosody acquisition. However, its hypotheses, principles, and predictions are still under investigation. Despite the abovementioned claims, on the phrasal level, intonation basically is the change in tune or pitch of the voice to communicate meaning based on varied linguistic contexts (Mennen, 2015). There are many linguistic uses for intonation in English, such as indicating the status of the phrase at the end of discourse. For instance, final falling intonation could suggest finality, while final rising intonation could indicate non-finality (Levis & Wichmann, 2015, p. 143).

Final intonation is one of the most important forms of intonation, where pitch moves from relatively high to relatively low (rising-falling, HL) or the opposite direction (falling-rising, LH) (Levis, 2018). In English, three aspects of intonation strongly influence intelligibility and comprehensibility: relative prominence, tune, and pitch range. Levis (2018) also postulated that “tunes are critical to how speakers construct discourse and communicate connections between and among utterances” (p. 175). A wide range of important meanings in English are related to variation in intonation, which affects intelligibility and comprehensibility.

Dissimilar to the patterns of final falling and rising intonation in English, Arabic tends to reduce intonation to a low fall at the end of phrases and sentences, particularly in reading

aloud (Swan & Smith, 2001, p. 199). This might be interpreted by native speakers of English as a lack of full and correct meaning completion. Another difference is the “sustained pitch” in Arabic, with pitch being steady on syllables, while pitch in English intonation tends to waver more (Kenworthy, 1987, p. 126). Mastery of L2 intonation thus represents a challenge for Saudi EFL learners, who could produce patterns different from English, potentially reducing intelligibility and comprehensibility.

### **3.4 Intelligibility and Comprehensibility in L2 Pronunciation**

Although L2 pronunciation has received more attention in education and research, it remains uncertain which L2 pronunciation constructs should be defined, conceptualised, and prioritised in instruction to make spoken communication more accessible. Munro and Derwing (1995, 1999, 2020) <sup>6</sup> defined and operationalised the constructs of intelligibility, comprehensibility, and accentedness with empirical evidence. Levis (2018) indicated that “Intelligibility is widely agreed to be the most important goal for spoken language development in a second language – both for listening and speaking – no matter the context of communication” (p. 15). Munro and Derwing (1999) defined intelligibility “as the extent to which a speaker’s message is actually understood by a listener” (p. 289). In other words, intelligibility denotes to what extent the speaker’s message matches the listener’s understanding. Levis (2018) defined comprehensibility “as the amount of work that listeners need to do in understanding a speaker” (p. 17).

Although intelligibility and comprehensibility are closely related, they are not interchangeable (Derwing & Munro, 2015). A speaker might be difficult to comprehend while

---

<sup>6</sup> The original work was produced in 1995. In 1999, it was reprinted in a special reprint of the decade’s best papers in the *Journal of Language Learning*. Munro and Derwing revisited their 1995 work in 2020 by commenting on their original methods and findings. For instance, statistical modelling analyses were employed, and the neglected or misunderstood parts of the original work were reconsidered.

still producing fully intelligible speech. Table 3.3 demonstrates the results of possible intelligibility and comprehensibility combinations from Derwing and Munro (2015, p. 6).

Another essential issue about intelligibility and comprehensibility is the way they are operationalised. Although there have been serious attempts to clarify their operationalisation, there is no consensus in the field of L2 pronunciation about how each speech parameter should be assessed (Kang & Ginther, 2018).

Table 3.3: Possible Intelligibility and Comprehensibility Combinations

Intelligibility	Comprehensibility	Results
High	High	Utterance is fully understood; little effort is required.
High	Low	Utterance is fully understood; great effort is required.
Low	Low	Utterance is not fully understood; great effort is required.
Low	High	Utterance is not fully understood; however, the listener has the false impression of having easily determined the speaker's intended meaning.

Thomson (2018) provided a detailed account of intelligibility and comprehensibility measurements. For intelligibility, Thomson summarised a number of studies that operationalised intelligibility in different ways, such as reaction times to the occurrence of intermittent, listener transcription scores for True/False sentences, two-way forced choice identification tasks, 10-way forced choice English vowel identification tasks, and listener transcriptions of sentences read aloud.

For comprehensibility, there has been a general consensus to use scalar ratings by following the approach of Munro and Derwing's 1995 study (e.g., Isaacs, 2008; Isaacs & Thomson, 2013; Kennedy & Trofimovich, 2008; Saito et al., 2016). However, a number of concerns may arise over the scalar ratings to assess how much effort the listener must make to understand a message, such as the size of the scale.

Although L2 pronunciation research is rapidly growing, more empirical research is needed as a variety of issues remain, including how to define and operationalise intelligibility and comprehensibility. As such, Thomson (2018) stated that "greater consistency in how

researchers define and operationalise these terms is necessary” (p. 13). Munro and Derwing (2015) stated that “although a number of studies of L2 speech intelligibility have appeared in leading journals in recent years, relatively little attention has been directed to establishing the validity and reliability of intelligibility measures” (p. 381). They suggested a number of approaches to assess L2 intelligibility, such as word count, sentence verification, cloze and dictation, content summaries, and comprehension questions. Some of these approaches might not be useful to quantify intelligibility without considering the nature of speech stimuli used.

Another issue is how intelligibility and comprehensibility are interrelated (Levis, 2005; Levis, 2020a). Although intelligibility must be prioritised, this does not mean comprehensibility is unimportant in successful communication (Munro & Derwing, 2015). Therefore, it is important to measure intelligibility and comprehensibility together.

### **3.5 The Efficacy of L2 Pronunciation Instruction**

#### **3.5.1 Intelligibility and Nativeness Principles**

Despite calls to focus on comfortable intelligibility, the early goal of pronunciation instruction was to make L2 speech similar to a native speaker (Abercrombie, 1949; Derwing & Munro, 2015). Furthermore, no explicit theoretical principles and guidelines or pedagogical implications and practices were proposed to promote intelligibility without imitating the model of native speakers. A turning point was when Levis (2005) argued the two conflicting principles (intelligibility and nativeness). The intelligibility principle refers to an L2 speaker’s need to be understood (Levis, 2020b). Intelligibility does not deny the effects of a strong foreign accent, but oral communication can be achieved regardless, as there is no direct link between accent and understanding. Although some segmental and suprasegmental errors can impair comprehensibility, comfortable intelligibility is still achievable. According to Levis (2018), greater priority in L2 pronunciation should be given to segmental and suprasegmental aspects

that could impede intelligibility. On the other hand, the nativeness principle asserts it is desirable to achieve native-like pronunciation, including among adult foreign-language learners (Levis, 2005). Although this was the primary aim of pronunciation instruction before the 1960s, it has become less important because it is unrealistic for many reasons, including biological constraints (Levis, 2005, p. 370).

With respect to ESL contexts, Murphy and Baker (2015) reviewed the four waves of ESL teaching pronunciation spanning 150 years: precursors (1850s to 1880s), the reform movement (1880s to early 1990s), the communicative approach, and empirical research (mid-1990s to the present). In the first three waves, the intelligibility principle was not the primary goal, which has changed in the fourth wave. These empirical studies can be categorised into three macro areas: what ESL segmental and suprasegmental features are important to teach, how to teach these features, and teachers and students' beliefs about pronunciation instruction. These themes have indicated that attaining comfortable speech intelligibility is the ultimate goal of pronunciation instruction.

However, it is unclear whether the intelligibility principle is achievable in an EFL context. Thus, there is a need for more empirical investigation of the effectiveness of pronunciation teaching in a Saudi EFL context based on what segmental and suprasegmental features should be taught and how to teach them to promote intelligibility and comprehensibility.

### **3.5.2 Effect of Instruction on Segmentals and Suprasegmentals**

Although Arabic and English could share the same phonetic and phonological patterns for certain features, such as stress, the ways they use them could differ. Thus, Saudi L2 English speakers often struggle with pronouncing English suprasegmentals, especially word stress, rhythm, and intonation, even when these features exist in Arabic phonology (Swan & Smith, 2001). In addition, many empirical studies have examined how L2 pronunciation instruction

can make L2 learners' speech more intelligible and comprehensible by targeting segmental and suprasegmental features. Although some studies were implemented in a short period of time (e.g., less than three weeks), the findings revealed instruction resulted in significant improvements to segmental and suprasegmental features (Derwing, 2018; Levis, 2018). This section reviews the potential benefits of L2 pronunciation instruction based on the length, focus, type, outcomes, and assessment of instruction, as these have been the core elements of empirical research (Thomson & Derwing, 2015). In addition, it is important to discuss whether it is beneficial to teach segmentals and suprasegmentals together or separately and which should take priority in the face of limited time and other constraints (Levis, 2018).

Saito (2012) reviewed 15 quasi-experimental studies to explore whether pronunciation instruction helped develop L2 pronunciation of segmentals and suprasegmentals. Three variables were used to code the studies. The first was the focus of instruction (segmental or suprasegmental). Saito examined five segmental and seven suprasegmental studies, while three studies focused on both. The second variable was type of instruction: focus-on-form (FonF) or focus-on-formS (FonFS); six studies were coded as FonF, 10 were FonFS, and one study used both. The difference between these types is that teachers in FonF might put some effort into controlled and communicative contexts, while teachers in FonFS provide only drills and activities for controlled contexts. The third variable was type of outcome measure: controlled constructed responses (CR) or free constructed responses (FR). Fourteen studies were classified as CR as they used word-, sentence-, and paragraph-reading tasks, while five studies were FR as they adopted picture description and a summary of brief lectures in a prepared topic (four studies used two types of measurements).

Saito (2012) found pronunciation instruction was effective in all segmental and suprasegmental studies except two, which were implemented in a very short time (15 to 30

minutes). Thus, length of instruction was crucial, with longer instruction showing more improvement. Regarding focus of instruction, all segmental studies displayed significant pronunciation improvement at a controlled level, and only one study showed improvement at a spontaneous level. All seven suprasegmental studies showed improvement on a controlled level, and one study measured on a free level demonstrated significant improvement. Regarding type of instruction, all six FonF studies revealed significant improvement at a controlled level and two at a spontaneous level (Saito, 2012). For FonFS instruction, eight studies measured and showed improvement at a controlled level, while three included measurements at a spontaneous level, but not all of them showed improvement. Regarding type of outcome measures, all 13 studies that used controlled constructed responses revealed improvement, but only two out of five that used free constructed responses displayed improvement. However, most of the 15 studies were conducted in English-speaking environments, no delayed posttest phase was implemented, and the participants were not beginning-level learners. In addition, it was unclear whether the ultimate goal of instruction in most of these studies was to attain intelligible and comprehensible pronunciation.

Thomson and Derwing (2015) surveyed 75 empirical investigations of L2 pronunciation instruction's effect in relation to participant demographics (e.g., L1 background, age, and language proficiency), theoretical parameters (e.g., nativeness or intelligibility), scope of instruction (e.g., segmentals or suprasegmentals), input of instruction (e.g., classroom or computer-assisted pronunciation teaching [CAPT]), duration of instruction (e.g., one semester or a few weeks), and description of assessment (e.g., speech tasks, control and comparison groups, and speech measurements).

The results showed that 74% of the 75 studies examined EFL and ESL learners, and these were primarily carried out in contexts where English had a major representation and

students could use English outside class (Thomson & Derwing, 2015). Most of these 75 studies recruited participants with different L1 backgrounds, making it difficult to determine which pronunciation features were important for specific groups based on their L1. Regarding theoretical parameters, 63% of studies assessed nativeness by looking at voice onset time (VOT) and pitch contours, 24% measured intelligibility and comprehensibility based on Thomson and Derwing's (2015) classification, and 13% assessed both. Thus, most of the reviewed L2 pronunciation research had no explicit theoretical foundation to determine which principle was important to measure.

With respect to scope of instruction, 53% of the 75 studies focused on segmentals, 23% focused on suprasegmentals, and 24% explored both (Thomson & Derwing, 2015). The results revealed no consistency in segment choice or focus of instruction. In addition, 61% of studies had traditional classroom-based instruction, while 39% had CAPT-based instruction. Duration of instruction varied dramatically, but classroom studies were generally conducted across more days or weeks with longer periods of instruction. Regarding task assessment, 73% of studies included a reading-aloud task, and 24% had a wordlist task. In addition, 79% recruited raters to assess the elicited L2 speech productions, and 21% used acoustic measures, such as VOT and formants (Thomson & Derwing, 2015). Finally, 60% of studies had a control group, and 40% had a comparison group. The findings of previous empirical studies must be treated with caution, however, as many factors could play a role in pronunciation, such as learner L1 background and length of instruction. Overall, some English segmental and suprasegmental features require more time and attention to enhance intelligibility and comprehensibility. Thus, instruction should target segmental and suprasegmental pronunciation that significantly impedes intelligibility and comprehensibility.

A long-standing question is whether it is more beneficial to teach segmental and suprasegmental features separately or together. Zielinski (2015) found both could increase intelligibility and comprehensibility and suggested there was no strong evidence to support teaching one over the other. Wang (2020) reviewed previous empirical research and found this debate has been overshadowed by other factors, such as L2 learners' needs and L1 background. Similarly, Celce-Murcia et al. (2010) believed "pronunciation instruction is moving away from the segmental/suprasegmental debate and towards a more balanced view" and that modern "curriculum thus seeks to identify the most important aspects of both the suprasegmentals and segmentals and integrate them appropriately in courses that meet the needs of any given group of learners" (p. 11). In other words, the priorities of teaching pronunciation should be set to maximise the benefits of instruction.

Munro and Derwing (2015) proposed several priorities to promote global intelligibility and comprehensibility. These included focusing on local phonological structures, having empirical evidence, emphasising pronunciation problems that do not resolve on their own, focusing on the most problematic errors among the majority of students, and assessing students individually. Levis (2016) suggested such priorities could be based on achieving comfortable intelligibility and comprehensibility, Jenkins' lingua franca core principle, and Brown's functional load principle. It is unclear, however, how to set priorities in relation to specific factors, such as learning context, L1 background, and language proficiency. Thus, studies "need to examine principles that help teachers decide on priorities based on context" (Levis, 2018, p. 53).

### 3.5.3 Effect of L2 Pronunciation Instruction over Time

Despite calls for using a delayed posttest phase to explore the long-term effects of L2 pronunciation instruction (Derwing & Munro, 2015), a great deal of empirical research has not implemented such a test (Levis, 2018, p. 81), which the present study employed.

Zhang and Yuan (2020) used a pretest, posttest, and delayed posttest across four months with 90 intermediate EFL learners at a Chinese university divided into a segmental focus group, a suprasegmental focus group, and no focus group in three intact classes. The delayed posttest was conducted three weeks after the intervention. The suprasegmental group maintained pronunciation gains over time for comprehensibility at the spontaneous level. A positive long-term effect of pronunciation instruction was also found by Lee et al. (2020). Participants were 115 low-intermediate students at a Japanese university who were categorised into five groups: syllabic perception, syllabic production, phonemic perception, phonemic production, and a control group. The duration between the pretest and delayed posttest was four weeks. Although the period of instruction was relatively short, the results revealed significant gains from the pretest to the delayed posttest. However, these empirical studies failed to demonstrate whether there were significant differences between the immediate and delayed posttests, as the baseline of comparison was only the pretest.

Couper (2006) explored the long- and short-term effects of two-week pronunciation instruction involving immigrant learners of English in New Zealand. A 99.9% average error rate in the pretest declined to 7.5% in the delayed posttest. Although the instruction was effective over time, the average error rate of the immediate posttest (5.5%) was slightly lower than the delayed posttest. French et al. (2020) examined the long-term effects of intensive pronunciation instruction on improving comprehensibility, enhancing fluency, and reducing foreign accentedness. The first group of participants ( $N = 42$ ) received intensive ESL instruction,

while the second ( $N = 39$ ) received non-intensive instruction. The comprehensibility and fluency gains from intensive instruction lasted longer.

Algethami (2017) examined the long-term effect of explicit phonetic instruction during a 15-week semester on producing English /p/. The experimental group showed no improvement over the control group. However, this study focused on only one segmental feature, and the sample size was small. Similarly, Ruellot (2011) examined the long-term influence of visual feedback on the two French vowels /u, y/ among 14 intermediate-level learners of French. The results revealed no significant improvement over time. However, the delayed posttest was conducted a week after instruction, which was short, and it was difficult to determine the influence over time.

Previous research has not shown reliable evidence for pronunciation gains over time with beginning-level learners in the Saudi EFL context. In addition, Thomson and Derwing (2015) stressed that “Delayed post-tests are necessary to measure whether instruction results in ongoing improvement relative to control groups” (p. 336). Thus, the lack of such tests in empirical investigations is a major challenge for L2 pronunciation research in general.

### **3.5.4 Effect of L2 Pronunciation Instruction on Beginning-Level Learners**

Another central issue in L2 pronunciation research and education is when to start incorporating pronunciation into the classroom. Much recent research has called for teaching pronunciation in the early stages. For instance, Zielinski and Yates (2014) suggested that pronunciation instruction could increase the confidence of beginning-level learners and that researchers and teachers have perceived this process as more difficult than it needs to be (p. 57).

In support of pronunciation instruction from the beginning of L2 learning, Derwing and Munro (2015) hypothesised a window of maximal opportunity, a period in which L2 learners show the most growth in L2 speech production and perception. This period seems to be in the

early stage of L2 learning, as the greatest phonetic and phonological shifts occur in their first six months of massive exposure to an L2. In a review of the state of the art of teaching pronunciation, Pennington (2021) noted that, despite a rapidly growing body of L2 pronunciation research, most empirical studies have been carried out on advanced students, with little attention given to beginners. Derwing (2018) recommended teaching pronunciation regardless of proficiency level:

Ultimately, the issue of when to start pronunciation instruction is tied to the needs of L2 speakers, regardless of where they are in their L2 learning trajectory. If their speech causes communication breakdowns, then they should be able to access pronunciation training that will result in improved communicative success. (p. 327)

Hence, the core consideration when offering pronunciation instruction is whether the learners need help with pronunciation and whether their pronunciation impedes communication. This consideration is applicable at all levels of proficiency.

Pronunciation difficulties could obstruct various aspects of development in the early stages of L2 learning, even when speakers have spent years in an English-speaking environment (Derwing, 2018). For instance, Derwing et al. (2014) found that L2 speakers who spent 19 years in Canada were not intelligible or comprehensible in their workplace. Once the participants received formal instruction, they became more understandable, and their perception achieved significant improvement. However, their foreign accent remained strong, and no improvements occurred in their speech fluency.

Springall (2002) conducted an action research study to explore whether pronunciation activities would help 21 beginning-level learners from nine language backgrounds. Beginners were comfortable with the new pronunciation concepts and terms and found the pronunciation activities useful, applying them in spontaneous speech. To explore beginning-level learners' perceptions of the influence of their pronunciation and interaction in English, Zielinski (2012) interviewed 26 Australian migrants (14 of them beginners). Almost 93% thought their

pronunciation difficulties made it harder for them to be understood. Furthermore, only beginners (23.1%) felt their pronunciation difficulties made them less confident and unwilling to communicate in English.

Despite its importance, there has been a lack of empirical work on the influence of pronunciation instruction on beginning-level learners in a Saudi EFL context. As a result, it is unclear whether such instruction could enhance intelligibility and comprehensibility, what learners think about their pronunciation difficulties, and how those difficulties could reduce their self-confidence. The present study sought to bridge this gap by implementing a pronunciation intervention for these learners followed by a questionnaire to explore their thoughts about the intervention.

### **3.6 Relationship of L2 Motivation to Intelligible and Comprehensible Speech**

Despite the consensus in the literature that adult L2 learners can acquire new sounds, additional factors could also affect the development of oral performance (Hansen Edwards, 2017; Moyer, 2014). One would be the motivation to learn and attitudes toward the L2 (Celce-Murcia et al., 2010, p. 279). According to Dörnyei and Ushioda (2011) “motivation theories [seek] to explain nothing less than why humans think and behave as they do” (p. 4). Dörnyei (2005) claimed motivation is the most basic force causing someone to begin and continue learning an L2 and is a component in other factors as well. In contrast, other factors, such as teachers or materials, cannot cause an unmotivated person to learn a language. Since the late 1950s, motivation theory in L2 acquisition has been theoretically argued, empirically tested, and pedagogically described. This research has had three phases. The first was the social psychological period (1959–1990), mainly influenced by Gardner and others in Canada. The second was the cognitive-oriented period (during the 1990s), primarily conceptualised by cognitive theories and educational psychology. The third is the process-oriented period (since

2000); characterised by the works of Dörnyei and others in Europe, this period seeks to examine motivation phenomena from a dynamic perspective.

As the first well-known attempt to conceptualise theoretical principles of L2 motivation, Gardner proposed an integrative and instrumental framework in 1985 (Dörnyei, 2005). The “integrativeness” of motivation is defined as how L2 learners show an explicit desire to be a member of the target language culture and community. Another major principle is the “instrumentality” of motivation, as it is related to practical, ongoing learning goals. Based on Gardner’s social-psychological theory, Gardner (1985) proposed an attitude motivation test battery, a questionnaire that measures a wide range of integrative and instrumental motivation components. Dörnyei (2005) claimed it has shown good psychometric characteristics to operationalise the main components of integrative and instrumental motives (p. 70).

However, Gardner’s theory failed to provide adequate explanations for critical issues in L2 motivation theory (Ryan, 2006; Ushioda, 2006). For instance, the principles of integrative and instrumental motivations were not clear from a classroom perspective. This is likely because the theory was conceived to provide theoretical explanations for the coexistence of two language communities (English and French) in Canada, a naturalistic setting. Another weakness is that it did not adequately explain the outcomes of L2 learners’ motivation in EFL settings from an integrative motivation standpoint. As Ryan (2006) noted, “what was true for the bilingual environment of Canada was not necessarily true for EFL learners lacking meaningful access to the target community and therefore the information necessary to develop a true desire to integrate” (p. 35).

Due to a lack of consistent evidence for Gardner’s theory, many have called for expanding it beyond an integrative-instrumental framework to include an education-centred approach that would be consistent with L2 teachers’ views and the outcomes of educational

psychological research (Dörnyei, 1994). Crookes and Schmidt (1991) claimed Gardner's theory did not account for other factors, such as the connection between motivation in L2 acquisition research and motivation as defined in other fields (e.g., cognitive and neuroscience studies). In addition, the theory failed to offer pedagogical implications of empirical findings. An initial attempt to expand the theory was made by Oxford and Shearin (1994) by conceptualising the cognitive development and its potential link to L2 learners' motivation. However, the influence of this attempt was limited because of a lack of empirical work supporting their theoretical arguments.

To examine L2 motivation learning from different angles (e.g., cognitive psychology), Dörnyei (2005, 2009) proposed a new model, the L2 Motivational Self System, based on two theories from cognitive psychology: possible selves theory (Markus & Nurius, 1986) and self-discrepancy theory (Higgins, 1987). The theory of possible selves seeks to dynamically connect one's self thoughts and self actions, while the theory of self-discrepancy examines self-images (actual self, ideal self, and ought-to self) to minimise the discrepancy between one's actual and ideal or ought-to self.

The core assumption of the L2 Motivational Self System is that L2 learners' explicit vision of their future selves shapes their behaviour toward an L2 (e.g., how often they use it) and their learning achievements (Dörnyei, 2005). The model has been conceptualised based on three dimensions: ideal L2 self, ought-to L2 self, and L2 learning experience. Ideal L2 self refers to the self-image of an ideal L2 user that a learner seeks to be. In other words, L2 learners are not satisfied with their current proficiency and aim to increase their linguistic competence. Ought-to L2 self is when L2 learners think they should improve their proficiency to meet specific expectations and avoid negative consequences. L2 learning experience refers to

learners' experiences and attitudes toward the learning environment, including teaching materials, methods, and teacher and peer behaviour.

Although a growing body of research has explored the relationship between L2 motivation and pronunciation development, there remains a lack of empirical evidence from different theoretical, empirical, and pedagogical perspectives. For instance, internal (motivation and attitudes) and external (experience and amount of L2 input and use) variables have not been a core interest of L2 motivation research on attaining intelligible and comprehensible speech. Instead, motivation has "been typically treated as a moderator variable measured using a few simple statements" (Trofimovich et al., 2015, p. 359). In addition, the L2 Motivational Self System does not explain all L2 motivation issues, and further theoretical, empirical, and practical investigations are needed to test theoretical assumptions, including the relationship between L2 pronunciation development and L2 motivation learning theory (Dörnyei, 2005).

In an early work on the relationship between motivation and L2 speech development, Moyer (1999) investigated the motivation of 24 advanced German learners, who pursued their graduate studies in German. Motivation was a strong factor for learners to achieve personal and professional goals. However, all learners were heavily accented in German, just over half stated it was important for them to have a native-like accent in German, and half indicated that being understood by native speakers was a realistic goal.

Saito et al. (2017) explored the impact of learner motivation on L2 pronunciation development in an EFL context. Forty first-year Japanese students were recruited to complete a questionnaire linked to their spontaneous speech development in terms of comprehensibility and accentedness as elicited from a picture description task at the beginning and end of one semester. Comprehensibility significantly improved among learners who were strongly motivated to study English to develop their future career as a long-term goal.

In another empirical study, Huensch and Thompson (2017) examined the relationship between learner attitudes and L2 pronunciation performance. The 195 participants completed an online survey in three parts (ideal, ought-to, and anti-ought-to selves). Learner attitudes showed they had no interest in sounding native-like but that pronunciation was still important to them. There was a positive relationship between positive attitudes toward pronunciation and the ideal self as well as between the desire to improve pronunciation and the anti-ought-to self.

Few studies have examined the L2 Motivational Self System in ESL/EFL contexts. For instance, Nagle (2018) examined the relationship between the development of L2 pronunciation and the dimensions of the system. Participants were 26 English-speaking learners of Spanish, who were asked to complete a picture description task five times in three semesters (nearly one academic year). Mixed-effects models revealed comprehensibility and accentedness improved over time. The dimension of intended effort (meaning the readiness of learners to put more effort into improving their English pronunciation) significantly reduced their foreign accent.

Measuring two dimensions of the L2 Motivational Self System (ideal L2 self and ought-to L2 self), Saito et al. (2018) carried out a cross-sectional longitudinal study to explore the development of comprehensibility among 108 high-school students in an EFL classroom based on motivation, emotion, and experience profiles. Students who frequently used English with positive emotions were directly influenced, which could improve comprehensibility over the long term.

With respect to the Saudi EFL context, most empirical studies have investigated L2 motivation as a predictor of achievement in a course or language based on the L2 Motivational Self System. For instance, Moskovsky et al. (2016) examined the relationship between learners' L2 motivational selves and L2 achievement. The study recruited 360 university students to complete a questionnaire consisting of items from each dimension of Dörnyei's theory. They

were also asked to take the writing and reading sections of an IELTS test to explore the relationship between English level and motivation. The results revealed no relationship between language proficiency and the L2 Motivational Self System. However, there was a significant relationship between writing scores and ought-to self.

Altalib (2019) examined Saudi university students' motivation in English for Specific Purposes (ESP) and English for General Purposes (EGP) courses. Over 4,000 participants were recruited to complete a questionnaire that adopted the components of the L2 Motivational Self System. Their motivation was found to be significantly related to the English courses they attended. In addition, the ESP group scored higher than the EGP group in ideal L2 self and showed more positive attitudes toward their L2 learning experience. However, no significant difference was found between the groups in terms of ought-to self.

Alshehri (2009) adopted the components of the L2 Motivational Self System to probe the impact of visual learning styles on students' ideal L2 selves. The results indicated a significant correlation between the effort students were willing to make to learn an L2 and their ideal L2 selves. Furthermore, a visual learning style was significantly related to ideal L2 self. However, participants' language proficiency levels were unclear. This made it impossible to determine whether a particular level was more motivated due to a significant relationship between learning style and the components of L2 self.

Despite broad agreement on the importance of motivation as a predictor of L2 pronunciation development (e.g., Celce-Murcia et al., 2010), there remains a lack of research on its role in making pronunciation more intelligible and comprehensible among Saudi EFL beginning-level learners. Furthermore, the relationship between the L2 Motivational Self System and L2 pronunciation development has been ignored by empirical L2 motivation studies in that context. Instead, most previous studies on the relationship between motivation and L2

pronunciation development have been conducted in English-speaking contexts (e.g., Huensch & Thompson, 2017).

In addition, numerous ESL studies have examined accent reduction rather than intelligibility and comprehensibility and did not take place in a classroom (e.g., Moyer, 1999). Participants in these studies were typically more proficient, making it unclear to what extent L2 motivation affected beginners. More empirical research is thus needed in the classroom. Finally, it is unclear which dimensions of the L2 Motivational Self System have a greater effect on L2 learners' oral performance at the beginning level in various EFL/ESL contexts.

### **3.7 Learner Perceptions of Pronunciation Instruction Effectiveness**

Research has seldom examined L2 learners' opinions about their pronunciation, what pronunciation challenges make them less intelligible and comprehensible, explicit pronunciation instruction, and non-pronunciation factors (e.g., learning environment, materials, methods, teachers, and motivation) in relation to intelligible and comprehensible speech. Eliciting these perceptions would help stakeholders improve pronunciation instruction design, set priorities, and maximise the effectiveness of instruction (e.g., Celce-Murcia et al., 2010; O'Brien, 2021).

Derwing (2003) explored how 100 intermediate ESL learners (who were all immigrants) perceived their pronunciation problems and the potential negative outcomes of having a foreign accent. Over 50% believed or strongly believed in the importance of improving pronunciation to avoid communication problems. In addition, the majority agreed their strong foreign accent should be reduced to integrate into the community.

In another empirical study, Scales et al. (2006) recruited 37 ESL learners and 10 American university students to listen to a passage read in four different English accents. Participants were asked to identify accents and determine their preferences for each accent.

Later, 11 learners were interviewed about their perspectives on the accents. Of these, 62% said their pronunciation goal in English was to attain native-like speech, but only 29% could identify an American accent.

At a U.S. university, Levis (2015) interviewed 12 advanced ESL learners about social issues linked to pronunciation instruction. They believed their accent needed to be like the models around them, such as a native English-speaking teacher. However, it was not clear whether learners would have the same views if they learned English pronunciation in an EFL environment.

Alameen (2014) explored L2 learners' perceptions of audio-visual and audio-only pronunciation instruction. Over half had positive perceptions about the training, finding it effective and helpful in linking speech production and perception. However, participants were not asked whether instruction helped make their speech more intelligible and comprehensible or sound like a native speaker of English.

To the best of my knowledge, no previous empirical study in a Saudi EFL classroom had explored learners' views about the pronunciation instruction they received to increase their speech intelligibility and comprehensibility. Furthermore, no previous research had investigated beginning-level EFL learners' perceptions about a variety of issues in pronunciation instruction. The present study sought to address this gap by asking such learners about the effectiveness of pronunciation instruction, teaching materials, the teacher's role, and the classroom.

### **3.8 Native-Speaker Perceptions of Non-Native Speakers**

There have been many calls to use native speakers to improve spoken communication with non-native speakers (Zielinski, 2008). This is because successful L2 spoken communication is the responsibility of non-native speakers and native listeners alike (Munro &

Derwing, 2015, p. 388). One way to make native speakers more involved is to elicit their perceptions about potential pronunciation challenges that could make non-native speakers less intelligible and comprehensible. However, much of the previous L2 speech perception and production research has focused on the role of native speakers to assess the success of L2 communication without asking them beyond their perceptual assessments of non-native speakers in order to obtain more in-depth understanding of the non-native speakers' pronunciation difficulties (Gordon & Darcy, 2016).

A growing body of empirical research is exploring the perceptions of native-English-speaking listeners in places where English is dominant. For instance, Hayes-Harb and Hacking (2015) carried out a mixed-methods study that asked 10 native English listeners to assess the accentedness of five Bosnian L2 speakers of English and five L1 speakers. This was followed by a semi-structured interview exploring the reasons for their assessments. Native speakers might rely on a complex set of attitudes and expectations to evaluate non-native speakers' pronunciation beyond a direct speech analysis.

Gordon and Darcy (2016) recruited 12 native English speakers to rate and comment on the pronunciation development of three groups of high-intermediate ESL learners (12 in each group) at a U.S. university. In the qualitative phase, native English listeners held similar views about which segmental and suprasegmental features made the non-native speakers' pronunciation more comprehensible. However, it was unclear whether the same comments would be made if the native-English listeners had assessed non-native speakers' intelligibility and comprehensibility.

In a study seeking to increase the role of L1 listeners in L2 spoken communication, Zielinski (2008) examined what segmental and suprasegmental patterns of L2 speakers might prevent intelligibility. L1 English listeners were found to rely on the same native processing

strategies to assess L2 speakers. Thus, the researcher asserted that native listeners needed to adjust their listening skills to non-native speech. However, the sample was limited to three non-native speakers and three native listeners.

Hayes-Harb and Watzinger-Tharp (2012) investigated how native-German listeners perceived the pronunciation of native-German speakers and native-English-speaking learners of German in terms of intelligibility and accentedness. Despite significant differences in the native listeners' qualitative data, the different perceptual assessments of non-native pronunciation did not necessarily reflect the beliefs of native listeners. In other words, different beliefs did not necessarily lead to different judgements of non-native pronunciation. Despite high inter-rater reliability of accent ratings and intelligibility judgements, the proficiency level of L2 learners was not homogenous.

Research is needed on native speakers' perceptions of non-native speakers in EFL contexts, such as Saudi Arabia. No previous research could be found that asked native listeners about the speech intelligibility and comprehensibility of beginning-level learners in ESL or EFL contexts. Hence, the present study sought to address this gap by asking native-English-speaking raters, once they completed the perceptual assessment task, about their perceptions of Saudi EFL learners' pronunciation. The results could clarify the pronunciation challenges preventing these learners from attaining more intelligible and comprehensible speech in English.

### **3.9 Research Questions**

To address gaps in the L2 pronunciation literature, the present study set out to answer the following research questions:

1. Does explicit instruction for segmentals increase the intelligibility and comprehensibility of Saudi beginning-level EFL learners more than non-explicit pronunciation instruction?

2. Does explicit instruction for suprasegmentals increase the intelligibility and comprehensibility of Saudi beginning-level EFL learners more than non-explicit pronunciation instruction?
3. Do Saudi beginning-level EFL learners retain pronunciation gains over time?
4. Is L2 motivation related to the development of more intelligible and comprehensible pronunciation among Saudi beginning-level EFL learners?
5. How do Saudi beginning-level EFL learners perceive the effectiveness of explicit pronunciation instruction?
6. How do native English-speaking raters perceive the pronunciation of Saudi beginning-level EFL learners?

### **3.10 Chapter Summary**

This chapter reviewed the literature on L2 pronunciation and presented the theoretical framework of the study. A descriptive overview of English segmentals and suprasegmentals identified how they are problematic for Saudi EFL learners. Intelligibility and comprehensibility are key concepts in L2 pronunciation research, and effective L2 pronunciation instruction is needed to help beginners produce more accurate segmentals and suprasegmentals. L2 motivation is another essential consideration because it could be related to L2 pronunciation development. Learners' perceptions about the effect of L2 pronunciation instruction can be used to improve instruction, while native English raters' perspectives can help identify learner challenges. The next chapter presents the research methodology.

## **Chapter 4: Methodology**

### **4.1 Introduction**

This chapter describes the methods of the study, including the design, participants, setting, instruments, pronunciation intervention framework and associated materials, data collection, pilot study, raters, internal and external validity, operationalisation of intelligibility and comprehensibility of segmental and suprasegmental features, and data analysis.

### **4.2 Research Design**

This study adopted a quasi-experimental pre-, post-, and delayed posttest design to see whether the speech intelligibility and comprehensibility of Saudi EFL learners could be improved through explicit instruction on segmental and suprasegmental pronunciation. Participants were divided into two experimental (segmental and suprasegmental) groups and a control group. In the first phase, a pretest measured the dependent variables (intelligibility and comprehensibility scores) prior to implementing the explicit pronunciation intervention. In the second phase, an immediate posttest determined whether the dependent variables displayed intelligibility and comprehensibility improvements as an immediate effect of the treatment. In the third phase, a delayed posttest identified whether the dependent variables showed intelligibility and comprehensibility improvements over a longer period of time.

To conceptualise the framework of the various types of experimental designs under the umbrella of quantitative research, it is essential that experimental studies deliberately manipulate the independent variables to determine to what extent the dependent variables can be influenced (Mackey & Gass, 2015, p. 189). Experimental designs can be classified based on the ability to control potential risks and maximise internal validity (Ary et al., 2019). For instance, a true experimental design is characterised by full control of confounding variables, randomisation, and a control group to increase internal validity.

In contrast, a pre-experimental design does not randomly assign participants and lacks a control group. This might lead to potential risks and the inability to control for confounding variables (cf. Ary et al., 2019). Although the quasi-experimental design does not randomly assign participants, it does employ useful strategies to control for confounding variables. Among these is assigning a control group and using intact classes. Therefore, the quasi-experimental design comprises the same standards as the true experimental design except it does not allow for randomly assigning participants. For a detailed methodological description of experimental designs, see Ary et al. (2019, pp. 247–271).

The current study took place in an EFL classroom in Saudi Arabia. A quasi-experimental design was more feasible because it was impossible to randomly assign participants as they were already assigned to intact groups (i.e., their classes). Consequently, the researcher would not have been able to manipulate and assign participants into random treatment groups.

Mackey and Gass (2015) pointed out that when participant randomisation is unfeasible, intact classes can be used to enhance the ecological validity of classroom research. Dörnyei (2007) also indicated that “because of the practical constraints, working with ‘*non-equivalent groups*’ has become an accepted research methodology in field studies where randomisation is impossible or impractical” (p. 117). Hence, the primary goal of the quasi-experimental design is to avoid unfeasibility and impracticability of random assignment in educational settings as that could lessen the effect of making a comparison between the assigned groups. Furthermore, Ary et al. (2019) postulated that the quasi-experimental design could help reach more reliable conclusions, although it is not possible to have maximum control of confounding variables (p. 260). In a similar vein, Ortega and Iberri-Shea (2005) claimed that empirical research that

employed a quasi-experimental design in recent years showed considerable improvement, including longitudinal investigations of L2 instructional effectiveness.

To enhance the quasi-experimental design adopted in the current study, two issues were taken into account. First, students were not able to self-select themselves to be in the treatment group. Second, differences between the two experimental groups and control group in the pretest phase were minimised to ensure the treatment phase was more effective (cf. Dörnyei, 2007). One of most practical techniques is to apply matching between participants to minimise differences and potential risks. Therefore, the current study was careful to achieve matching between participants in terms of their demographics, language background, and English proficiency, as all the participants were beginning-level learners.

In addition, the research design included four questionnaires. A demographic and language background questionnaire elicited data such as age and experience learning English. A second questionnaire with questions about motivational and attitudinal characteristics in L2 learning was used to obtain a deeper understanding of whether individual variability (e.g., motivation) was likely to play a role in achieving more improvement in L2 speech intelligibility and comprehensibility among Saudi EFL learners. A third questionnaire was used after the pronunciation intervention to explore the experimental groups' perspectives on the effectiveness of the treatment in general and with specific segmental and suprasegmental patterns. Finally, a questionnaire elicited the raters' perceptions about the speech intelligibility and comprehensibility of learners throughout the three tasks (pre-, post-, and delayed posttests).

There are a number of advantages to questionnaires in L2 research. For one, they can help elicit longitudinal information from participants in a brief span of time (cf. Mackey & Gass, 2015). This information can be compared among a number of participants. In addition, questionnaires are more economical and practical than individual interviews, and they can be

administered in different ways, such as via email, phone, or in person. This allows researchers to have a higher degree of flexibility in data collection. Based on how questions are structured, questionnaires can provide profound quantitative and qualitative insights and are suitable for a broad range of research (Mackey & Gass, 2015, pp. 102–105).

### **4.3 Participants and Setting**

Participants consisted of 45 Saudi beginning-level adult EFL learners, who completed a battery of pre-, immediate post-, and delayed posttests as well as questionnaires. The participants were assigned into three intact groups. The first was the first experimental group (segmental group), which consisted of 15 participants who received explicit pronunciation instruction on segmentals. The second experimental group (suprasegmental group) consisted of 15 participants who received explicit pronunciation instruction on suprasegmentals. A control group consisted of 15 participants who received non-explicit instruction.

All participants were male students between the ages of 18 and 20 ( $M = 18.69$ ,  $SD = .668$ ), were native speakers of the Northern Saudi dialect of Arabic, and were born and raised in Al-Jouf Province, where the study took place (see Table 4.1). With respect to university major, 23 participants were studying in the College of Medicine (51.1%) and 22 in the College of Engineering (48.9%). The only languages participants had studied were Arabic and English. They started learning English in elementary and intermediate school: 28 started when they were 12, five when they were 13, one when he was 14, and 11 at the age of 6. Thirty-three participants had studied English for eight years in public school, four had studied it for six years in public school and six months in private English institutions, and nine had studied it for nine years in private schools. The participants were taught by native English-speaking teachers, including Americans, British, Australians, New Zealanders, and Canadians. No participants had received pronunciation training or visited English-speaking countries. With respect to using English

outside of class, 35 reported using English via television, social media, mobile learning apps, websites, movies, private tutors, and games. They reported a variety of difficulties being understood by native English-speaking teachers; 15 faced difficulties pronouncing words and sentences, six had difficulties understanding their teachers, eight lacked accuracy in pronunciation, 10 had a lack of confidence, and six had a fear of making pronunciation mistakes.

Table 4.1: Participants' Language Background Information

Language Background	Number of Participants
Native language/dialect	- Arabic ( $n = 45$ ) - Northern Saudi dialect ( $n = 45$ )
University major	- Medicine ( $n = 23$ ) - Engineering ( $n = 22$ )
Languages they had studied or been exposed to	- Arabic ( $n = 45$ ) - English ( $n = 45$ )
Age and grade level at which they began studying English	- Age 12 (intermediate) ( $n = 28$ ) - Age 13 (intermediate) ( $n = 5$ ) - Age 14 (intermediate) ( $n = 1$ ) - Age 6 (elementary) ( $n = 11$ )
Years of English before enrolling in the English programme in the PYP	- Public school (7 years) ( $n = 33$ ) - Public school (6 years) / private institutions (6 months) ( $n = 4$ ) - Private school (9 years) ( $n = 8$ )
Previous exposure to English pronunciation training	No ( $n = 45$ )
Taught by native English-speakers	Yes ( $n = 45$ ) (Americans, British, Australians, New Zealanders, and Canadians)
Visiting English-speaking countries	No ( $n = 45$ )
Using English outside of class and reasons for doing so	- Yes ( $n = 35$ ) (TV, social media, mobile learning apps, websites, movies, private tutor, and games) - No ( $n = 10$ )
Difficulties being understood by native English-speaking teachers	- Difficulty pronouncing some words and sentences ( $n = 15$ ) - Difficulty understanding teachers ( $n = 6$ ) - Lack of accuracy in pronunciation ( $n = 8$ ) - Lack of confidence ( $n = 10$ ) - Fear of making pronunciation mistakes ( $n = 6$ )

Participants were recruited from the class of English 1 (ENGL01) in the Deanship of Common First Year (DCFY) at Al-Jouf University in northern Saudi Arabia. The primary purpose of this class was to improve learners' English listening, speaking, reading, writing, grammar, and vocabulary so they could communicate and interact in English. This class was

compulsory for students expected to pursue their undergraduate studies in English-medium instruction. The textbooks were part of a series: *New Headway Elementary, Student Book*. Students' English proficiency was evaluated by the deanship based on the English Language Placement Test (ELPT), which divided proficiency into beginner, intermediate, and advanced.

Participant selection involved convenience sampling, i.e., selecting participants who were easily accessible with characteristics that met the study purposes (Creswell & Creswell, 2017). Dörnyei (2007) pointed out that “the most common sample type in L2 research is the convenience or opportunity sample, where an important criterion of sample selection is the convenience of the researcher” (p. 98).

To maximise sample validity and apply the matching technique to minimise potential risks of the quasi-experimental design, participants had to meet the following criteria. First, since the study sought to investigate novice EFL learners, the participants had not been exposed to English in English-speaking environments. Second, they had no previous pronunciation training. Third, to control the impact of dialectal variation in Saudi Arabia, the study was only carried out on native speakers of the Northern Saudi dialect of Arabic, and all were born and raised in the same region. Fourth, participants had no hearing deficits. Fifth, only male undergraduate students participated, as it would have been difficult for the male researcher to recruit female students due to a lack of access to their campus. Sixth, participants who did not complete all tests and questionnaires or failed to attend all intervention sessions were excluded. Thus, four participants in the segmental group, three in the suprasegmental group, and four in the control group were excluded. This resulted in each group having 15 participants.

#### **4.4 Instruments**

Three instruments were used to answer the research questions: questionnaires, segmental stimuli, and suprasegmental stimuli.

#### 4.4.1 Questionnaires

The first questionnaire collected demographic information and participants' language background. The second explored their perceptions about the effectiveness of pronunciation instruction. The third was used to see which motivational profiles were related to the development of intelligibility and comprehensibility. The fourth elicited raters' perceptions of the audio-recordings they assessed and how they found the L2 speech intelligibility and comprehensibility of participants. The questionnaires for students were carefully translated into Arabic by a certified experienced English-Arabic translator and verified by another. In addition, the translated questionnaires were piloted to remove any ambiguity, as explained in Section 4.9.

The first questionnaire consisted of two parts (see Appendix A). The first included seven questions asking about nicknames, gender, age, birthplace, where they grew up, university major, and whether they had any speaking or hearing difficulties. The second part included nine questions about first dialect/language and number of languages they had studied or been exposed to. Furthermore, they were asked what age they had first been exposed to English and whether they had studied English before enrolling in the English programme in the PYP. Participants were also asked whether they had received any type of pronunciation training, whether they had taken English phonology and phonetics courses, and whether the trainer or teacher was a native speaker of English. The questionnaire asked whether they had been in English-speaking countries and for their purpose and length of stay. As non-classroom exposure to English is highly important, participants were asked whether they were exposed to English outside of the class, such as watching TV and using English on mobile apps. Finally, they were asked to explain what difficulties they had faced understanding their English teachers.

The questionnaire was developed and inspired by Marian et al. (2007), who collected demographic and language proficiency data from bilingual participants, as this strategy has

been widely used in various disciplines of social sciences, including linguistics (Kaushanskaya et al., 2019). Some questions were modified to obtain demographic and language information pertinent to this study. The questions also ensured participants were homogenous to control for confounding variables that could cause individual differences that might play a role in improving their pronunciation beyond the influence of the pronunciation treatment.

The second questionnaire investigated the relationship between L2 motivation and attitude toward learning English in attaining intelligible and comprehensible L2 pronunciation (see Appendix B). The questionnaire was based on the L2 motivational self-system and attitudes toward learning English, which were proposed in a recent dynamic approach to explore to what extent L2 motivation and attitudes were related to L2 learning, including L2 phonology acquisition (see Dörnyei, 2009). Four themes were addressed. The first explored learners' intended efforts to learn English. The second investigated the ideal image (ideal L2 self) L2 learners would have in the future. The third examined how L2 learners should work hard and take responsibility to meet others' expectations. The fourth theme investigated attitudes toward learning English. Each theme consisted of five items presented in random order and measured on a 6-point Likert scale ranging from "strongly disagree" to "strongly agree". The items were adopted and modified from Taguchi et al. (2009).

The third questionnaire explored participants' perceived satisfaction with the pronunciation intervention and its outcomes (see Appendix C). The questionnaire consisted of seven close-ended questions on a 4-point Likert scale ranging from 1 "Strongly disagree" to 4 "Strongly agree" and three open-ended questions. The open- and close-ended questions were adopted and modified from Alameen (2014). The close-ended questions elicited participants' perceptions about whether the treatment was effective, realistic, and practical. In addition, they were asked whether they had observed any pronunciation improvements after the treatment, the

appropriateness of the length of the treatment, and to what extent the treatment was suitable for their proficiency level. Two questions were about the helpfulness of the materials and the activities used in the treatment. With respect to the open-ended questions, the first and second were about what participants found to be the most and least beneficial components in the treatment. The third question asked what participants would change to improve the treatment.

The fourth questionnaire was divided into two parts (see Appendix D). The first collected the demographic information, language background, and teaching experience of the raters. More specifically, the raters were asked about their gender, age, place of birth, native accent, how long they had taught, whether they had taught in an EFL context, linguistics and English courses they had taught, and whether they had had any contact with Arabic-speaking learners of English. The second part explored raters' perceptions about the audio-recordings they assessed. They were also asked to explain which specific phonological features might prevent Saudi EFL learners from attaining intelligible and comprehensible speech in English. The questionnaire was adopted and modified from Gordon and Darcy (2016).

#### **4.4.2 Segmental Stimuli**

This section describes the segmental stimuli used to elicit participants' L2 speech intelligibility and comprehensibility. To date, there has been no consensus on which segmentals should be prioritised, as they cause tremendous difficulties for English learners and must be taught at different levels of proficiency and in different contexts (Derwing & Munro, 2015). This complexity stems from a variety of reasons, such as English learners having different L1 backgrounds. With respect to the current study and because the participants were homogenous in their demographics and language background, segmental stimulus items were selected based on the most problematic English consonants for Saudi EFL learners, as these sounds do not have counterparts in the Arabic sound system in terms of a phonetic contrastive analysis. Thus,

these sounds might lead to loss of speech intelligibility and comprehensibility when communicating in English (see Ahmad & Muhiburrahman, 2013; Ahmad & Nazim, 2014; Alfallaj, 2013; Algethami, 2016, 2017; Alharbi, 2015; Ali, 2015; Alsofyani & Algethami, 2017; Hameed & Aslam, 2015; Shah & Al-Bargi, 2016). The segmentals used in the current study were /p, v, ʃ, ŋ/.

To measure L2 speech intelligibility, segmentals were introduced in high-frequency monosyllabic words to be consistent with participants' beginner English proficiency level. In addition, these segmentals were presented in words found in the beginner and elementary levels of the *New Headway* textbooks that participants were studying in the PYP at Al-Jouf University when the study took place. The words were chosen from content words, such as main verbs, nouns, adjectives, and adverbs, because they are more important to reach a deeper lexical meaning (Gilbert, 2008). In addition, segmentals were introduced in initial and final positions, except for /ŋ/, which only appeared in final position because no English words use /ŋ/ in initial position (Rogerson-Revell, 2018, p. 103). It is worth noting that medial position was excluded to avoid using disyllabic words to be more consistent with participants' proficiency level. Furthermore, the number of segmental stimuli was determined based on participants' beginner status, as it would be unrealistic and counterproductive to test and teach participants all the problematic segmentals for Arabic learners of English. In addition, the intervention was limited to three weeks, so it was not practical within the scope of this study to test and teach all problematic English segmentals.

Furthermore, the segmental stimuli were presented in minimal pairs that were identical except for one sound. These were used because sounds in minimal pairs are predicted to be the most difficult to differentiate in terms of L2 speech production and perception (Best & Tyler, 2007). To give a practical example, because of the lack of phonetic realisation of the differences

between the Arabic and English sound systems, the Saudi EFL learners might tend to replace English phonemes not found in Arabic phonology with their closest equivalents, such as replacing /p/ (a sound absent from learners' dialect of Arabic) with Arabic /b/ (Algethami, 2016). Each contrast was presented in six minimal pairs of English words as the target sound appeared in three words in initial position and three in final position. Positional effects were used to see whether they led to problems with L2 pronunciation. These words were only used in the battery of tests and not in the pronunciation instruction phase to avoid participants having any potential familiarity with them, which could make it harder to identify pronunciation improvements. The same stimulus words were used in the pre-, immediate post-, and delayed posttests as it was impossible to find a different set of real words with the same characteristics for each test (i.e., being common for beginners, being content words, and having the same number of syllables).

A month transpired between each testing phase to reduce the possible influence of participants' memory. In addition, to avoid making participants familiar with the stimulus words, the words were presented in a random order and in a carrier sentence by following Hayes-Harb et al.'s (2008) approach, such as "*I like to say **pack** more than **back***". This helped ensure participants did not realise the target segments were not the same or even minimal pairs. In addition, the stimulus words were piloted to ensure they were aligned with the English proficiency of participants so they could produce the sentences as naturalistically as possible. The list of the words used as segmental stimuli to measure speech intelligibility is given in Appendix E.

To measure the potential effect of the treatment on segmental patterns in L2 speech comprehensibility, the same problematic segmentals used to measure intelligibility were presented in real sentences, although the actual words in the intelligibility stimuli were not used to measure comprehensibility because they were isolated words. Furthermore, there is a long-

standing debate about the importance of measuring speech intelligibility and comprehensibility in different ways and by employing different sets of stimuli.<sup>7</sup> Therefore, a different set of stimuli was employed to provide valid, reliable data on comprehensibility.

In terms of linguistic complexity, a set of issues was considered in the stimulus sentences to measure comprehensibility. At the phonological level, for instance, the sentences only included the target problematic segmentals. At the morphosyntactic level, the sentences were designed to have appropriate grammar and vocabulary for the beginner level. At the lexical level, the target segmentals in the given sentences were very frequent content words for beginners based on the Word Frequency (n.d.) corpus. Furthermore, the sentences were piloted to ensure they were aligned with participants' proficiency level so that they could produce the sentences as naturalistically as possible. For the stimulus sentences used for measuring segmental speech comprehensibility, see Appendix E.

#### **4.4.3 Suprasegmental Stimuli**

This section describes the suprasegmental stimuli used in the battery of tests. Suprasegmental features (lexical stress, rhythm, and final intonation) for speech intelligibility and comprehensibility were selected because they would be more challenging for beginning-level Saudi EFL learners, as the phonological patterns of the selected English suprasegmental features differ from Arabic phonology according to the principles of contrastive phonology (Alfallaj, 2013; Altaha, 1995; Swan & Smith, 2001). These suprasegmental features were generally among the most important for beginners that could impede intelligibility and comprehensibility (Brinton, 2014; Celce-Murcia et al., 2010; Grant, 2014b). Another consideration was the appropriateness of these suprasegmentals for beginning-level learners and the amount of instruction employed to avoid overwhelming participants.

---

<sup>7</sup> For more about the measurements of intelligibility and comprehensibility, see Section 4.7.

With respect to lexical stress patterns, the study examined the stress placement in disyllabic noun-verb pairs. Such pairs of words have the same spelling but different stress placement due to being in different word classes (Celce-Murcia et al., 2010). In terms of rhythm, the study investigated how participants, after being taught, were aware of the rhythmic patterns in English sentences. In particular, they were taught to differentiate between content words (i.e., nouns, main verbs, adjectives, adverbs), which are usually stressed, and function words (e.g., articles, pronouns, prepositions), which are mostly unstressed. Final intonation patterns were examined to explore how participants dealt with the final falling pitch at the end of statements (indicating certainty) and the falling pitch at the end of questions (indicating uncertainty).

Several variables were considered when designing and selecting suprasegmental stimuli (Colantoni et al., 2015). First, it was essential to consider word frequency because less frequent words might not be appropriate to beginning-level learners. To avoid this issue, all suprasegmental stimuli were selected based on the Word Frequency (n.d.) corpus. Second, neighbourhood density was considered to measure how a set of words would be consistent in terms of their phonology, orthography, and semantics in order to avoid the effect of neighbourhood density on the processing of stimulus words and sentences.

Third, it was essential to control the effect of lexical category on word stress stimuli in terms of intelligibility and comprehensibility. Fourth, syntactic complexity was considered to avoid presenting suprasegmental stimuli in sentences with an overly complex structure for lower-proficiency learners. Fifth, the types of pragmatic knowledge were considered to be related to how beginners use and interact in English in a real and communicative way. Sixth, three linguists and two native speakers of English checked the suprasegmental stimuli to identify any problems or complexities at the lexical, phonological, syntactic, morphological,

and pragmatic levels. Finally, all suprasegmental stimuli were piloted to ensure they were aligned with the English proficiency of participants.

For lexical stress stimuli, the study adopted 10 pairs of isolated words (20 words total) with two different possible pronunciations based on whether they were a noun or a verb to measure intelligibility (Prator & Robinett, 1985). These more frequent disyllabic words were selected based on the proficiency level of participants. The words were only used as stimuli in a battery of tests to examine whether learners could place stress on the correct syllable in each pair. None of these words were used in the instruction phase to avoid making participants familiar with the words. These stimuli were presented in isolated words in carrier sentences, such as “*I like to say \_\_\_ more than \_\_\_*”. Each word in the pair was in one of the blank spaces, e.g., “*I like to say **record** more than **record***”. The list of words used as suprasegmental stimuli to measure L2 speech intelligibility is given in Appendix F. To measure the L2 speech comprehensibility of lexical stress, the same word pairs from the intelligibility test were presented in read-aloud sentences rather than isolated words. Ten sentences were used with five target words containing lexical stress as a noun and the other five as a verb. These words were presented in simple grammatical sentences. The sentences used in the lexical stress comprehensibility task are given in Appendix F.

For rhythmic stimuli, 10 sentences containing 24 stressed content words were presented in read-aloud sentences to measure both speech intelligibility and comprehensibility. The frequency of words used as stimuli and grammatical structures were in keeping with learners’ proficiency level. The sentences used as stimuli were adopted from Grant and Yu (2016), as shown in Appendix F. For final intonation stimuli, 20 stimulus sentences were used to measure both intelligibility and comprehensibility of final intonation, and they were presented in read-aloud sentences. The sentences were all structurally simple and contained common words that

could have different meanings based on falling or rising final intonation. The sentences are shown in Appendix F.

## **4.5 Pronunciation Intervention**

### **4.5.1 Pronunciation Intervention Framework**

This section describes the pronunciation intervention framework for the two experimental groups and the control group. The framework for the experimental groups followed a Listening and Awareness-Control-Practice-Extension (LCPE) sequence—an explicit, systematic, and communicative approach proposed for beginning-level learners (Zielinski & Yates, 2014). The adopted approach allowed students to advance through a sequence of four stages of development. The goal of the first stage was to increase their listening ability and awareness of the target features. The second stage aimed to help them control physical properties to correctly pronounce the target features. The third stage had them practise in a different set of structured contexts, including difficult ones, while the fourth stage helped them apply what they had learned in new uninstructed contexts. Table 4.2 shows the four stages of the explicit and systematic approach of pronunciation instruction, taken from Zielinski and Yates (2014, p. 66).

To enhance this intervention, different modalities and techniques were used to explicitly demonstrate the target features to students. Specifically, auditory, visual, and kinaesthetic modalities were used to advance participants through the four stages of development and avoid some target features going unnoticed. Furthermore, different types of techniques, such as stretching a rubber band, were applied to highlight the importance of the target features. Table 4.3 (reproduced from Zielinski & Yates, 2014, p. 69) provides examples of the modalities and techniques implemented in each stage.

Table 4.2: Goals of Each Stage of Development

Stage of Development	Aims
Listening and awareness	<b>To develop learners’ awareness of the target pronunciation feature and how it differs from the feature in the L1.</b> Learners need this exposure in order to discover the physical and perceptual aspects of the target English sound or pattern. At this stage learners might, for example, develop their ability to identify words that start with the target sound, words that have the same stress pattern, or words that are emphasised in a particular phrase.
Control	<b>To develop learners’ physical control over the pronunciation of the target feature.</b> At this stage learners might, for example, work on a particular sound at the beginning of words, the production of words with a particular stress pattern, or the emphasis of the appropriate word or phrase.
Practice	<b>To develop learners’ ability to produce the target feature in a range of different and increasingly difficult structured contexts.</b> For example, learners might start by practicing a pronunciation feature in single words and then progress to pronouncing that feature in short phrases and then longer sentences.
Extension	<b>To develop learners’ ability to apply their new skills in a range of contexts.</b> At this stage, learners might practice the target sound or pattern in somewhat less structured activities such as answering questions or participating in short scripted dialogues. They might then progress to using that feature in slightly more spontaneous classroom contexts (e.g., asking for directions, making appointments, or participating in everyday conversations).

Both experimental groups were given feedback in each developmental stage. As Zielinski and Yates (2014) stated, “it is crucial that we provide feedback for beginners sensitively and judiciously so that we don’t undermine their confidence” (p. 68). This feedback was designed to be realistic by targeting students’ needs since the ultimate goal was to attain intelligible and comprehensible speech in English. Feedback was also systematic and relevant to each stage of development. For instance, feedback to attain free-error pronunciation in the control stage would have been unrealistic. It was also useful to know when to provide immediate and delayed feedback based on stage of development and whether items were new or had been used in previous stages.

The control group received the non-explicit pronunciation intervention but engaged with the same materials as the experimental groups without going through the four systematic stages of development of the explicit intervention. Traditional techniques such as “listen and repeat” were used to broadly demonstrate pronunciation features and to deliver the materials

and communicative activities without explicitly teaching participants to be aware of difficulties that could make their pronunciation unintelligible or incomprehensible. Furthermore, no explicit feedback was provided to correct specific pronunciation errors. Overall, the non-explicit pronunciation intervention was considered effective as it had been implemented in a great deal of previous research to explore the potential effects of explicit instruction on L2 speech perception and production compared to non-explicit interventions (e.g., Gordon & Darcy, 2016).

Table 4.3: Using Modality to Teach Stress Patterns

Stage of Development	Technique	Modality and Examples
Listening and awareness	<b>When providing examples of target words</b> , students listen to the teacher and watch the accompanying movements. They observe the teacher using movements (e.g., clapping, snapping the fingers, tapping the desk, stretching a rubber band, taking a step, opening a fist, etc) to correspond with the production of stressed syllables in the words.	Visual/Auditory. For example, the teacher closes fist next to face and open fist when saying the stressed syllable in va-CA-tion
	Stress patterns of target words are represented using dots or Cuisenaire rods above the syllable.	Visual. For example: • 0 • to MA to
Control	Learners use movement to correspond with stressed syllables of the target words.	Visual/Kinaesthetic. For example: the learner stretches a rubber band and watches movement while saying the word and stressing the appropriate syllable in va-CA-tion
	Learners use dots or Cuisenaire rods above the syllable as a guide to the stress patterns of the target words as they say them aloud.	Visual: • 0 • va CA tion
Practice	Learners practice using the target words in context. At first, stressed syllables might be capitalised, marked with dots, or accompanied by movement (as described above), but cueing decreases as students' progress.	Auditory/Visual/kinaesthetic
Extension	Learners use words in less structured activities such as dialogues or role plays. The teacher uses various modalities to provide reminders and feedback to the learners if they have difficulty with particular stress patterns. Learners might also use various multimodality techniques to reinforce stress patterns for themselves as they practice words in different contexts.	Auditory/Visual/kinaesthetic

#### 4.5.2 Pronunciation Intervention Materials

Materials and activities in explicit and non-explicit instruction were adopted from the textbook *Well Said Intro: Pronunciation for Clear Communication* (Grant & Yu, 2016). This textbook is designed to help beginning-level learners of English improve pronunciation and ultimately attain intelligible and comprehensible speech. The second edition includes the following key features based on a growing body of research in L2 pronunciation:

- A clear, easy-to-follow course plan for teachers with a limited background in pronunciation instruction and/or limited time to plan lessons.
- Primary focus on the learner goal of improving intelligibility rather than reducing a foreign accent.
- Prioritised instruction, based on the listening/speaking needs of beginning to low-intermediate students as well as the speech features that contribute most to overall intelligibility.
- Increased attention to perception and focused listening, and specifically to helping students notice critical differences between their production and the more intelligible target form.
- Updated vocabulary and useful phrasal contexts as informed by corpus linguistics.
- A chapter organisation that spirals and recycles learning.
- A focus on both form (through structured and guided pronunciation practice) and meaning (through naturalistic speaking activities).
- End-of-chapter recorded tasks that promote targeted self-monitoring and provide for individualised feedback from the teacher.
- An online workbook with additional pronunciation/listening practice. (Grant & Yu, 2016, p. ii)

This textbook has been found to be effective in terms of its course content and organisation, course preparation, metacognitive rationale, activity modification, and additional resources (Lima, 2017; Zimmerman, 2018). Hence, it was aligned with the needs of the current study, as illustrated by Table 4.4 (reproduced from Grant & Yu, 2016, pp. vi-viii), which exhibits the scope and sequence of the chapters from *Well Said Intro* that were used in the pronunciation intervention to teach the target segmental and suprasegmental features.

Table 4.4: Textbook Summary

Feature	Part/ chapter	Objectives	Rules and practice	Communicative practice	Pronunciation log
Supra- segmentals	Part III (Ch. 7): Stress in nouns and verbs	<ul style="list-style-type: none"> <li>To make syllables in words sound stressed.</li> <li>To use simple guidelines to determine the stress in nouns and verbs.</li> </ul>	<ul style="list-style-type: none"> <li>Stress in two-syllable nouns.</li> <li>Stress in two-syllable verbs.</li> <li>Stress in two-syllable noun and verb pairs.</li> </ul>	Troubles with technology.	<ul style="list-style-type: none"> <li>Record and submit an announcement</li> <li>Self-monitor for stress in nouns and verbs.</li> </ul>
	Part IV (Ch. 9): Rhythm: Stressed words	<ul style="list-style-type: none"> <li>About English rhythm in phrases and sentences.</li> <li>Which words are stressed.</li> <li>Why words are stressed.</li> </ul>	<ul style="list-style-type: none"> <li>Content words and sentence stress.</li> <li>Structure words and sentence stress.</li> </ul>	Voicemail.	<ul style="list-style-type: none"> <li>Record and submit limericks.</li> <li>Self-monitor for rhythm of stressed words.</li> </ul>
	Part V (Ch. 13): Final intonation	<ul style="list-style-type: none"> <li>How the voice rises or falls at the end of sentences.</li> <li>About intonation patterns in statements and questions.</li> </ul>	<ul style="list-style-type: none"> <li>Falling intonation statements.</li> <li>Falling intonation: <i>Wh-</i> questions.</li> <li>Rising intonation: <i>Yes/no</i> questions.</li> </ul>	Roommates.	<ul style="list-style-type: none"> <li>Record and submit questions.</li> <li>Self-monitor for final intonation.</li> </ul>
Segmentals	Part I (Ch. 3): Consonant sounds	<ul style="list-style-type: none"> <li>An overview of four consonant sounds used in the current study.</li> <li>The four consonant sounds and their common symbols.</li> <li>The four consonant sounds and spelling.</li> </ul>	Various communicative practices for each target sound in the <i>Well Said Intro</i> Online Workbook		<ul style="list-style-type: none"> <li>Tying the knot: Who pays?</li> <li>Nine lives.</li> <li>Take a vacation!</li> <li>The best diet.</li> </ul>
	Part VI: Consonant sounds				

#### 4.6 Data Collection Procedures

This section describes the four phases of data collection (see Table 4.5). The first covers participant recruitment, the demographic and language background questionnaire, pretest recording, and the L2 motivation questionnaire. The second describes how teachers were prepared, the pronunciation intervention implementation, immediate posttest recording, and the post-intervention questionnaire. The third describes the delayed posttest recording. The fourth phase explains the perceptual assessment of the audio-recordings and the post-assessment questionnaire for the raters. The ethical aspects of the study procedures were reviewed and approved by the Australian National University (ANU) Human Research Ethics Committee (Protocol 2018/208).

Table 4.5: Data Collection Phases

	Phase I	Phase II	Phase III	Phase IV	
Participants	Demographic and language background questionnaire + pretest + motivation questionnaire (Week 1, 2)	Instruction (Week 3, 4, 5)	Immediate posttest + post-intervention questionnaire (Week 6, 7)	Delayed posttest (Week 11, 12)	Assessment of recordings + post-assessment questionnaire (10 weeks)
Exp Group 1	✓	Explicit	✓	✓	X
Exp Group 2	✓	Explicit	✓	✓	X
Control Group	✓	Non-explicit	✓	✓	X
Raters	X	X	X	X	✓

#### 4.6.1 Phase I: Recruitment, Background, Pretest, and Motivation Questionnaire

After obtaining approval from the Deanship of Common First Year (DCFY) at Al-Jouf University, a formal announcement was made to students taking English 1 (ENGL01) with the approval letter (see Appendix G). The data were collected at the beginning of the 2018 fall semester because participants would have had little to no exposure to English, reducing any non-experimental factors affecting their pronunciation. The first phase of data collection consisted of four main parts: recruitment and assignment of participants, the demographic and language background questionnaire, the pretest audio-recordings, and the L2 motivation questionnaire.

Since the first phase took place over the course of two weeks, the first day was dedicated to recruiting 45 participants and assigning them into two experimental groups and a control group. Each group included 15 participants. After this stage, participants were asked to carefully read the participant information sheet that was translated into Arabic (for the English version see Appendix H and for the Arabic version see Appendix I). It included a detailed description of the study and the purpose and length of each task. The participants' questions were explicitly answered, and they were assured all necessary steps would be taken to respect their privacy and confidentiality and that they could withdraw at any time. In addition, they could remove any

part of their responses they were not satisfied with. Subsequently, participants were asked to sign the consent form after carefully reading its content (for the English version see Appendix J and for the Arabic version see Appendix K). On the second day, the demographic and language background questionnaire was handed out to the class and took approximately 15 minutes to complete (for the Arabic version of the questionnaire, see Appendix L).

As three groups were assigned, each participant individually recorded the stimulus items in a sound-isolated booth in the English laboratory in the Deanship of the First Common Year at Al-Jouf University. Starting on the third day, each group completed the audio-recordings in two days with 15 minutes recorded from each participant. This process took six days for all three groups. A Zoom H5 Handy Recorder was used at a sample rate of 44.1 kHz, 16 bit. Each stimulus item was presented in a separate PowerPoint slide with a short pause between slides to avoid any corruption in the recording.

Stimulus items were randomly presented to ensure participants were not aware of the target problematic segmentals and suprasegmentals. Before the pretest recordings, participants were given five minutes to go through the slides and give a dry run with the recorder to avoid any influence from lack of comfort and were encouraged to relax. No feedback was given once they completed the pretest recordings.

The fourth part of the first phase of data collection was the L2 motivation questionnaire, which was handed out in class (for the Arabic version, see Appendix M). The questionnaire took about 15 minutes for each participant to complete.

#### **4.6.2 Phase II: Teacher Preparation, Intervention, Immediate Posttest, and Post-Intervention Questionnaire**

The second phase of data collection took place in Weeks 3–7 and was composed of four parts. The first was to recruit and prepare two teachers by following Gordon and Darcy's (2016)

approach to avoid teaching fatigue. The first teacher was an experienced native speaker of English, who was asked to deliver the pronunciation intervention to the two experimental groups. To enhance the effectiveness of this intervention, the teacher was recruited because he held a master's in Teaching English to Speakers of Other Languages (TESOL) and had seven years of experience in the Saudi EFL context. He also had adequate knowledge of the most problematic aspects of pronunciation preventing Saudi English speakers from being intelligible and comprehensible, as he received pedagogical and phonetic training about these problems.

The second teacher was recruited to deliver the non-explicit intervention to the control group. He was a native speaker of English with five years of experience teaching English in Saudi Arabia and had also received formal training on the pronunciation problems of Saudi EFL learners. Two teachers were recruited because one teacher could experience instructor fatigue or get confused about the objectives of each group as a result of delivering two different treatments to three groups in a short, intensive period of three weeks. The teachers were paid.

A week before the intervention, the researcher met with the teachers to provide the teaching framework, materials, lesson plans, and instructional guidelines for the intervention. The teachers were informed in detail about the nature of each session and the anticipated intervention outcomes. In addition, they were informed that the researcher would attend all pronunciation intervention sessions for the three groups to verify intervention fidelity, take notes, and check attendance. Since the resources were handed over to the teachers, they were asked to prepare before each session.

The second part of the second phase was the pronunciation intervention. The three learner groups received the intervention for 25 minutes at the end of class three days per week (Monday, Tuesday, and Wednesday) for three weeks. In the experimental groups, the first teacher followed the explicit approach (developmental stages), techniques, and modalities. In

the listening and awareness stage, they received explicit phonetic instruction and an analytical overview of the target feature to develop their awareness. For the control stage, they were taught to develop physical control over the pronunciation of the target feature.

For the practice stage, the experimental group developed the ability to produce the target feature in a range of increasingly difficult structured contexts. They started by practising a pronunciation feature in single words and progressed to short phrases and longer sentences. Different tasks were carried out, such as minimal-pair recognition and discrimination and analysis of stress and rhythm in short passages. In the extension stage, they were taught to apply what they had learned in the previous three stages to a range of unstructured or less structured contexts, such as answering questions or participating in short scripted dialogues. They then progressed to using that feature in a more spontaneous classroom context, such as asking for directions, making appointments, or participating in everyday conversations.

The second teacher followed the non-explicit approach with the control group. These participants only received an announcement that pronunciation would be practised. They were asked to listen to and repeat words, phrases, and sentences. Although all three groups had similar production tasks, such as pair and group discussions, roleplaying, and information-gap activities, the control group did not go through the experimental group's stages of development in the explicit instruction.

The third part of the second phase was the immediate posttest audio-recordings over a period of six days, with each group taking two days to complete the task. To minimise potential differences, the same procedures from the pretest were strictly followed for the immediate posttest. To avoid order effects influencing the results, the stimulus items were randomly ordered in a different way from how they had been presented in the pretest.

The fourth part of the second phase was the post-intervention questionnaire (for the

Arabic version, see Appendix N). It was distributed in class and took about 10 minutes to complete. After this stage, participants were informed that the last phase of data collection would be carried out in a month.

#### **4.6.3 Phase III: Delayed Posttest**

The third phase of data collection took place in Weeks 11 and 12, comprising the delayed posttest audio-recordings. This task was completed in six days, with two days assigned for each group, and followed the same procedures as the previous tests. To reduce any possible influence from order effects, the stimulus items were randomly ordered in a different way from how they had been presented in the previous tests.

#### **4.6.4 Phase IV: Perceptual Assessment of Recordings and Post-Assessment**

##### **Questionnaire**

Before the perceptual assessments, the target segmental and suprasegmental features were extracted and isolated using Praat (Boersma & Weenink, 2019). This software allowed the researcher to complete the process of extracting, cleaning, and preparing the data, such as removing pauses and carrier sentences. After this stage, the audio files were inserted into Qualtrics (Qualtrics, 2020). Because of the demanding rater selection criteria, it was impossible to ask all raters to complete assessments in the same place. Hence, Qualtrics with strict conditions was employed, as it is widely used in experimental linguistic research (e.g., Woods et al., 2015).

The audio-recordings of each feature were put in one block, with a total of nine blocks, as all recordings of all features could not be put in one block. Splitting the recordings into blocks was also an effective way to avoid fatigue as raters could not complete all assessments in one session. For easy reference, each recording in each block was coded. For instance, “E1P01T1S1” refers to Experimental Group 1 (E1), Participant 1 (P01), Pretest (T1), and Segmental 1 (S1),

whereas “E2P01T2Rhy” indicates Experimental Group 2 (E2), Participant 1 (P01), Immediate Posttest (T2), and Rhythm (Rhy). The codes were not seen by the raters.

The first block contained the demographic questionnaire items and the audio-recordings of segmental /p/ from both groups in all three tests. The second block comprised the recordings of segmental /v/, the third contained segmental /tʃ/, and the fourth had segmental /ŋ/ and the post-assessment questionnaire to elicit raters’ perceptions about what they had assessed in the audio-recordings of the four segmental features. Unlike the first four blocks that measured the intelligibility of segmental audio-recordings, the fifth included the segmental recordings aiming to assess speech comprehensibility and questions about how the raters assessed the comprehensibility of the segmentals.

Regarding suprasegmentals, the sixth block contained the audio-recordings of lexical stress in isolated words to measure intelligibility along with questions about those recordings. The seventh block consisted of recordings of lexical stress in sentences to measure comprehensibility in addition to related questions. The eighth block consisted of the recordings of rhythm to measure intelligibility and comprehensibility as well as related questions. Finally, the ninth block consisted of the final intonation audio-recordings to measure intelligibility and comprehensibility along with the questions about the measurements of these constructs. Table 4.6 illustrates how the audio-recordings of each feature were split into nine blocks in Qualtrics.

To increase the inter-rater reliability of assessments, the audio-recordings were randomly presented to avoid any influence from increased familiarity with voices. Furthermore, the raters were asked to listen only once, and if they were not sure, they were asked to guess. They were also not able to download the audio-recordings. To ensure no question or audio-recording was left unassessed or unanswered, all responses were compulsory.

Prior to posting the advertisement to recruit raters, the audio-recordings in each block,

instructions, questionnaire, time limit for each block, technical issues, and potential difficulties were piloted by a linguistics PhD student to explore how feasible the assessments appeared. No major risks or difficulties were noticed. Furthermore, raters were asked to have a trial audio file before the formal assessments to ensure the volume was appropriate, remove ambiguity, avoid technical difficulties, and ensure the raters understood the tasks.

Table 4.6: Division of Audio-Recordings into Blocks

Block	Feature	Questionnaire
1	/p/ (in isolated words)	Demographics
2	/v/ (in isolated words)	
3	/tʃ/ (in isolated words)	
4	/ŋ/ (in isolated words)	Questions about segmental audio-recordings for intelligibility
5	All four target segmental features (in sentences)	Questions about segmental audio-recordings for comprehensibility
6	Lexical stress (in isolated words)	Questions about lexical stress audio-recordings for intelligibility
7	Lexical stress (in sentences)	Questions about lexical stress audio-recordings for comprehensibility
8	Rhythm (in sentences)	Questions about rhythm audio-recordings for intelligibility and comprehensibility
9	Intonation (in sentences)	Questions about intonation audio-recordings for intelligibility and comprehensibility

An advertisement to participate in the study was distributed to the teachers' community through the researcher's social circle (see Appendix O). Ten experienced native English-speaking teachers were recruited to rate participants' audio-recorded tests. The raters were asked to carefully read the instructions (see Appendix P) along with the participant information sheet (see Appendix Q) and then sign the consent form (see Appendix R).

After this stage, the raters were given instructions for what to do before the assessments. For instance, they were told they could complete the task on their laptops at a time and place convenient for them with an emphasis on the quality of headphones and the quietness of the location. They were also told to take breaks as much as they wanted. Although each block took approximately 10 to 25 minutes to complete, the raters were asked to avoid closing the page before completing a block to keep the responses from being lost. Once the raters completed this

work, they emailed the researcher to arrange the payment. The perceptual assessments of segmental and suprasegmental features for intelligibility and comprehensibility and the questions about these assessments took 10 weeks to complete.

#### **4.7 Operationalisation of Speech Intelligibility and Comprehensibility Constructs**

There has been a long-standing debate about the relative reliability of measurements of intelligibility and comprehensibility in L2 speech. Thus, it remains unclear which measurement approach would be more reliable (Derwing & Munro, 2015; Kang et al., 2018; Thomson, 2017). Measuring comprehensibility would appear to be easier than intelligibility as a Likert-type scale can be employed to rate how easy or difficult a pronunciation is for native speakers to understand (Thomson, 2017). Table 4.7 demonstrates the operationalisation of speech intelligibility and comprehensibility constructs in the current study.

To measure the intelligibility of the segmental stimuli, a two-way forced choice identification task of minimal pairs was used. Raters were asked to identify which word they heard. This approach has been found very helpful to precisely measure listeners' ability to understand the intended message (Thomson, 2017). In addition, a two-way forced choice identification task has been used in a number of well-cited papers (e.g., Hayes-Harb et al., 2008; Sheppard et al., 2017). The isolated words were presented in random order to enhance inter-rater reliability of measurements. Raters were asked to guess if they were unsure.

To measure L2 speech intelligibility of lexical stress stimuli, a two-way forced choice identification task was used to measure the disyllabic noun and verb pairs showing lexical stress. The words were presented in random order to raters, who were asked to listen and choose between two options (noun or verb). This approach with lexical stress has been used in previous research and has been found more reliable for measuring L2 speech intelligibility (e.g., Frost, 2011; Lepage & Busà, 2014; Tremblay, 2009). Furthermore, Thomson (2017) pointed out that

“forced-choice tasks, in which listeners identify the sounds they perceive, are a more direct measure of intelligibility” (p. 8).

Table 4.7: Operationalisation of Constructs

Feature	Operationalisation of Construct	
Segmental/ Suprasegmental	Intelligibility	Comprehensibility
All four target segmental features (in isolated words)	Two-way forced choice identification of minimal pairs.	
All four target segmental features (in sentences)		Rating sentences read aloud from 1 (very difficult to understand) to 9 (very easy).
Lexical stress (in isolated words)	Two-way forced choice identification of noun and verb differentiated by stress placement.	
Lexical stress (in sentences)		Rating sentences read aloud from 1 (very difficult to understand) to 9 (very easy).
Rhythm (in sentences)	Transcription of sentences.	Rating sentences read aloud from 1 (very difficult to understand) to 9 (very easy).
Final intonation (in sentences)	Two-way forced choice identification of statement and question differentiated by final intonation.	Rating sentences read aloud from 1 (very difficult to understand) to 9 (very easy).

To measure the intelligibility of rhythm stimuli, a transcription task was used, in which raters transcribed the sentences they heard. A transcription approach was used because the goal of the rhythm stimuli was to see where participants placed stress on each word in each sentence in relation to the rhythm rules of English. To measure the intelligibility of the final intonation stimuli, a two-way forced choice identification task was used. The raters were asked to listen and choose whether they heard a sentence as a statement or a question by using the falling/rising pattern of final intonation as a cue.

On the other hand, the comprehensibility of the segmental, lexical stress, rhythm, and final intonation stimuli was measured using a 9-point Likert scale that rated sentences read aloud from 1 (very difficult to understand) to 9 (very easy to understand). This scale has been found to be valid and reliable in measuring comprehensibility in several studies (Derwing & Munro, 1995; Munro & Derwing, 2015). In addition, the 9-point Likert scale has been used

successfully in a large body of research, and the results have revealed high inter-reliability (Gordon & Darcy, 2016).

Finally, it is worth mentioning that a controlled task was used rather than spontaneous tasks because the participants were beginning-level learners with limited access to prosodic features that could help them be more fluent. Thus, spontaneous speech might have been overwhelming for beginners (e.g., Zielinski & Yates, 2014). However, this type of speech assessment could be addressed in higher proficiency levels (e.g., upper intermediate and advanced).

#### **4.8 Raters and Inter-Rater Reliability of the Audio-Recording Assessments**

As noted above, 10 experienced native English-speaking teachers were recruited to rate the audio-recordings. Their ages ranged from 39 to 51 ( $M = 45$ ,  $SD = 4.136$ ) and their teaching experience in English from 10 to 22 years ( $M = 15.40$ ,  $SD = 4.195$ ). All the raters said they had never taught native Arabic speakers of English. In addition, their teaching experience in EFL had not been in any Arabic-speaking countries but rather, for example, in Russia, Europe, China, Thailand, Japan, South Korea, India, and Vietnam. The pronunciation and linguistics courses they had taught included oral communication in English, academic English courses, pronunciation courses at various levels, integrated skills, and introduction to linguistics.

A great deal of research has argued that experienced raters could play a key role in obtaining more objective assessments and in-depth subjective reports, in particular if they have not had any contact with participants' L2 accent or native language (Gordon & Darcy, 2016). Therefore, to maximise inter-rater reliability of perceptual assessments, raters had to meet the following criteria: 1) they were native speakers of English, 2) they were teachers of English with at least five years of experience, 3) they had not taught Arabic speakers of English, 4) they

had not lived or taught in Arabic-speaking countries, and 5) they had an academic background in linguistics and language teaching.

Inter-rater reliability was calculated by computing an intra-class correlation coefficient (ICC) via R (R Core Team, 2019) using *irr* (Gamer et al., 2019). The ICC was selected because it is more appropriate in measuring internal consistency between multiple raters (Koo & Li, 2016; McGraw & Wong, 1996). The results revealed a strong level of reliability between raters, as the average ICC was 0.91 for segmental intelligibility, 0.94 for segmental comprehensibility, 0.87 for lexical stress intelligibility, 0.82 for lexical stress comprehensibility, 0.85 for rhythm intelligibility, 0.95 for rhythm comprehensibility, 0.86 for final intonation intelligibility, and 0.92 for final intonation comprehensibility. When the size of the ICC of internal consistency is above .90, the level of reliability is excellent and is good when above .80 (Cicchetti, 1994, p. 286). Therefore, the ratings for intelligibility and comprehensibility were highly reliable.

#### **4.9 Pilot Testing**

Mackey and Gass (2015) noted that “a pilot study is an important, if not essential, means of assessing the feasibility and usefulness of the data collection methods and making any necessary revisions before they are used with the research participants” (p. 52). Thus, as a crucial step prior to the main study, a number of issues were piloted. This was done to (1) test the feasibility of the quasi-experimental design in a Saudi EFL context; (2) assess whether the stimuli were appropriate for participants’ proficiency level; (3) explore whether the questionnaires had ambiguity or redundancy and improve questions, formats, and instructions that were translated into Arabic; and (4) identify and avoid potential challenges to data collection.

To pilot the feasibility of the quasi-experimental design, a one-day pronunciation treatment with a small sample was conducted in July 2018 during the summer semester at Al-

Jouf University. The sample was divided into an experimental group and a control group. Each group had two participants, who had the same characteristics as the participants in the main study. The experimental group received an explicit treatment on the segment /p/, while the control group received the non-explicit treatment on the same segment. The participants were asked to record the stimuli for /p/ before and after the treatment.

Pilot testing showed the recruitment technique and research design to be viable. The experimental group showed significant improvement over the control group, suggesting the stimuli for the segment /p/ in the treatment were effective.<sup>8</sup> Furthermore, the oral feedback from participants indicated the treatment and stimuli for /p/ were feasible and useful. The participants in the pilot test were excluded from the main study.

To assess the appropriateness of the stimuli to participants' English proficiency level, two participants in the pilot study were asked to record all segmental and suprasegmental stimuli once. This step ensured the items could be produced comfortably by beginning-level learners. The same two participants completed questionnaires to see whether the translation was clear, accurate, and easy to understand. Based on their feedback, some items in the L2 motivation questionnaire were modified. For instance, the participants were confused by the phrase "I can imagine..." in the items for the "ideal L2 self" theme. The participants assumed the phrase referred to a different meaning from its intended meaning. Some typos were also corrected.

To identify and avoid potential risks in the data collection procedures, the place and process of recording were piloted. The place was appropriate and no obstacles were observed

---

<sup>8</sup> The pilot data were assessed by one English-speaking rater to look for improvements after explicit instruction. No modelling statistical analyses were performed because this was beyond the goal of pilot testing the quasi-experimental design and stimulus items.

as it was a sound-isolated booth in the English laboratory in the Deanship of the First Common Year at Al-Jouf University, ensuring high recording quality.

In terms of the process of recording, the participants displayed a good understanding of what the pretest and posttest recording phases required. In addition, the time needed to complete the recordings and fill out the questionnaires were piloted to avoid participant fatigue. No major changes were required as the participants could complete the tasks without fatigue. Transferring the data from the recording device to the researcher's Mac laptop was also piloted to check for file corruption and ensure the data could be stored securely. The outcome of this step was positive, and no risk was observed.

#### **4.10 Validity of the Study**

There are numerous systematic ways to measure the effectiveness of a study in terms of validity and inter-rater reliability. Ary et al. (2019) stated that the relationship between the variables of a study can be taken as an inference to determine whether the variables are valid. Therefore, any potential risks must be addressed to enhance validity. There are several types of measuring validity. One is internal validity, which seeks to verify the extent of the effect of variable A (the treatment) on variable B (the outcome) by controlling any extraneous variables influencing the outcomes. Another is external validity, which explores whether the results of the cause-effect relationship can be generalised to other participants, contexts, and measurements.

A number of issues were considered to enhance the internal validity of this study. One potential risk to internal validity was a lack of randomisation. Three steps were taken to minimise this risk. The first was to assign a control group, and the second was to use an intact class. Using the intact class was also beneficial to avoid selection bias that could be a risk to internal validity as a random assignment of participants was not feasible. The third step was to

apply a matching technique among participants to control the effect of individual differences among participants. Therefore, participants had identical or nearly identical demographic and linguistic characteristics. This step helped validate the sample by making participant demographic and language background relatively homogenous.

Another step to enhance internal validity was to use the same segmental and suprasegmental stimuli throughout the three tests, albeit in a random order in each test (cf. Ary et al., 2019). This helped ensure that no changes in instruments could play a negative role in the outcomes. As a further step, the pilot study offered an opportunity to internally validate the main study, including the instrumentation.

Experimental mortality, which refers to the loss of participants in different stages of an experiment, was a serious risk to internal validity (Ary et al., 2019). In the second week of the pronunciation treatment, four participants in the segmental group, three in the suprasegmental group, and four in the control group were absent. To reduce this risk, 60 participants were recruited and divided evenly across the three groups. Hence, at least 15 participants in each group attended all sessions, which was the target number for each group.

On the other hand, controlling or reducing risks to external validity is not as straightforward as risks to internal validity (Ary et al., 2019). Two risks to external validity were taken into consideration. The first was the subject effect, as participants might have negative feelings that developed during the study. For instance, some participants might feel nervous about individually recording stimulus words and sentences in the lab. To control this risk, the researcher explained the lack of negative consequences from their performance in the study.

The second risk to external validity was to select a small sample that represented the same language background. Nevertheless, limitations of the research problem, design,

objectives, and setting required a small sample with the same language background. To reinforce the ecological validity of the study and generalisability of the results, participants received the pronunciation treatment in a classroom (a real-life context) that could be similar to a different population in other contexts.

#### **4.11 Statistical Analysis Procedures**

To prepare the data for analysis, the segmental and suprasegmental datasets for intelligibility and comprehensibility were exported from Qualtrics into Excel spreadsheets in .csv format. The datasets were converted from wide to long format to be ready for analysis in R. It is important to clarify that the primary aim of this study was not to compare the segmental and suprasegmental features, as each feature was tested separately.

As the SLA field still widely uses certain statistical tests (e.g., ANOVA and *t*-tests), in recent years more calls have been made to adopt mixed-effects models. For instance, Cunnings (2012) strongly encouraged researchers to use statistical modelling. He stated that “mixed-effects models provide the second language researcher with a powerful tool for the analysis of a variety of types of second language acquisition data” (p. 369).

As another advantage, Linck and Cunnings (2015) pointed out that “Mixed-effects models also easily allow for the inclusion of multiple participant-level and stimulus-level independent variables in a single analysis, potentially offering a fruitful way of examining how individual differences may affect L2 acquisition” (p. 186). As a result, the current study ran a series of mixed-effects models to predict the significance between and within groups across the tests along with the descriptive analysis via means and standard deviations (SD).

For the intelligibility of the segments /p/, /v/, and /ʃ/, three generalised linear mixed-effects models were fit in R (R Core Team, 2019) using lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2017). The perceived intelligibility of a correct response was added as a

dependent variable. The fixed effects in the full model comprised a three-way interaction between group (experimental, control), time (pretest, posttests), and the position of the target segment in each stimulus item (onset, coda). These factors served as predictors of response correctness, while participant, rater, and word were random effects. The formula for the full model for the intelligibility of the target segment was  $\text{Correctness} \sim \text{Group} * \text{Time} * \text{Position} + (1|\text{Word}) + (1|\text{Participant}) + (1|\text{Rater})$ . If an interaction or fixed effect was non-significant, the model was pruned until only significant predictors remained ( $p < 0.05$ ). For the intelligibility of the segment /ŋ/, a generalised linear mixed-effects model was fit in R by applying the same formula as the full models of the other segments with the same modelling procedures, except that the predictor of position was not included in the full model.

For the intelligibility of lexical stress, a generalised linear mixed-effects model was fit in R using the same formula as the segmental full models and following the same modelling procedures but without position as a fixed effect in the full model. With respect to the intelligibility of rhythm, a linear mixed-effects model was fit in R with intelligibility correctness as a dependent variable. The fixed effects in the full model included a three-way interaction between group (experimental, control), time (pretest, posttests), and sentence type (A, B). Participant, sentence, target word, and rater were random effects. The formula for the full model for rhythm intelligibility was  $\text{Score} \sim \text{Group} * \text{Time} * \text{Sentence type} + (1|\text{Target.word}) + (1|\text{Sentence}) + (1|\text{Participant}) + (1|\text{Rater})$ . For intelligibility of final intonation, a generalised linear mixed-effects model was fit in R using the same formula and modelling procedures as before but without position as a fixed effect and with the random effect “word” replaced with “sentence” in the full model. Therefore, the full model for final intonation was formulated as follows:  $\text{Correctness} \sim \text{Group} * \text{Time} + (1|\text{Participant}) + (1|\text{Sentence}) + (1|\text{Rater})$ .

For the comprehensibility of segmentals, lexical stress, rhythm, and final intonation, a set of linear mixed-effects models was fit in R. The comprehensibility score served as a dependent variable, and the fixed effects in the full model included a two-way interaction between group (experimental, control) and time (pretest, posttests) as predictors. Participant and rater were random effects. The formula for the full model for perceived comprehensibility was  $\text{Score} \sim \text{Group} * \text{Time} + (1|\text{Participant}) + (1|\text{Rater})$ . The sjPlot package was used for visualising the model output of segmentals and suprasegmentals for intelligibility and comprehensibility (Lüdtke, 2018). Table 4.8 provides a summary of statistical modelling analyses that were used based on the independent and dependent variables of segmental and suprasegmental features and the nature of the L2 speech intelligibility and comprehensibility constructs.

For motivation, a linear mixed-effects model was fit in R. The motivation score served as a dependent variable. The fixed effects in the full model included a two-way interaction between motivational dimensions (intended effort, ideal and ought-to L2 self, and attitudes toward learning English) and group (segmental, suprasegmental, and control). Participant and question were random effects. The formula for the full model was  $\text{Score} \sim \text{Group} * \text{Motivational Dimension} + (1|\text{Participant}) + (1|\text{Question})$ .

For the post-intervention questionnaire, a linear mixed-effects model was fit in R. The score served as a dependent variable. The fixed effects in the full model included a two-way interaction between theme (intervention satisfaction, intervention outcomes) and group (segmental, suprasegmental). Participant and question were random effects. The formula for the full model of the post-intervention questionnaire was  $\text{Score} \sim \text{Group} * \text{Theme} + (1|\text{Participant}) + (1|\text{Question})$ . Table 4.9 demonstrates the statistical analysis procedures for motivation and the post-intervention questionnaire.

Table 4.8: Modelling Analysis Procedures for Segmental and Suprasegmental Features

Construct	Dependent Variable	Independent Variable	Modelling
Intel.	Intelligibility correctness (categorical variable)	<b>For /p, v, tʃ/:</b> Time, group, and position (fixed effects). Three-way interaction between groups (experimental, control), time (pre-, immediate post-, delayed posttests), and position (onset, coda). Participant, word, and rater (random effects). <b>For /ŋ/:</b> Time and group (fixed effects). Two-way interaction between group and time. Participant, word, and rater (random effects).	Generalised linear mixed effects
Intel.	Intelligibility correctness (categorical variable)	<b>For lexical stress:</b> Time and group (fixed effects). Participant, word, and rater (random effects). <b>For final intonation:</b> Time and group (fixed effects). Participant, sentence, and rater (random effects). <b>For both:</b> Two-way interaction between group and time.	Generalised linear mixed effects
	Intelligibility score (continuous variable)	<b>For rhythm:</b> Time, group, and sentence type (fixed effects). Three two-way interactions between group, time, and sentence type. Participant, sentence, stressed word, and rater (random effects).	Linear mixed effects
Com.	Comprehensibility score (continuous variable)	<b>For all segmentals and suprasegmentals:</b> Time and group (fixed effects). Two-way interaction between group and time. Participant and rater (random effects).	Linear mixed effects

Content analysis was used with the qualitative data elicited from three open-ended questions addressing Research Questions 5 and 6 (Dörnyei, 2010). In practical terms, the key points were identified in participants' corresponding responses and highlighted. Excerpts from the responses were taken to support the line of argument.

Table 4.9: Modelling Analysis Procedures for Motivation and Post-Intervention Questionnaire

Research Question	Dependent Variable	Independent Variable	Modelling
Motivation	Motivation score (continuous variable)	Group and motivational dimensions (fixed effects). Two-way interaction between group and motivational dimensions. Participant and question (random effects).	Linear mixed effects
Post-intervention	Intervention score (continuous variable)	Group and theme intervention (fixed effects). Two-way interaction between group and motivational dimensions. Participant and question (random effects).	Linear mixed effects

#### **4.12 Chapter Summary**

This chapter described the methods used in the current study. This included a detailed description of the research design, participants, sampling, instruments, pronunciation intervention framework and materials, and the four phases of data collection procedures. The chapter also explained the pilot study carried out before the main study. Rater demographic information was provided, along with the inter-rater reliability of the audio-recording assessments. In addition, the chapter addressed the validity of the study, operationalisation of intelligibility and comprehensibility, and procedures for a set of mixed effects models. The results of the current study are presented in the next chapter.

## Chapter 5: Results

### 5.1 Introduction

This chapter presents a detailed report of the results in nine sections. The second and third report descriptive statistical and mixed-effects modelling results for the intelligibility and comprehensibility of segmentals. The fourth and fifth present those results for suprasegmentals, while the sixth does so for motivation and attitudes in L2 learning. The seventh section presents learners' perceptions about the effectiveness of the explicit pronunciation intervention, and the eighth presents how raters perceived learners' pronunciation. The last section summarises the chapter.

### 5.2 Intelligibility of Segmentals

This section reports the descriptive and modelling results of the four segments in terms of intelligibility.

#### 5.2.1 The Segment /p/

##### 5.2.1.1 Descriptive Results for /p/ Intelligibility

The participants produced the segment /p/ 540 times in six-word pairs in each test battery. These occurrences were assessed by 10 raters. Thus, 30 participants  $\times$  3 tests  $\times$  6 words  $\times$  10 raters produced 5,400 occurrences. Table 5.1 shows a summary of the descriptive results for the intelligibility of /p/.

Table 5.1: Descriptive Statistical Summary for /p/ Intelligibility

Group	Test Response					
	Pretest		Immediate Posttest		Delayed Posttest	
	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
Experimental	642 (71.3%)	258 (28.7%)	233 (25.9%)	667 (74.1%)	627 (69.7%)	273 (30.3%)
Control	578 (64.2%)	322 (35.8%)	365 (40.6%)	535 (59.4%)	543 (60.3%)	357 (39.7%)

As seen in Table 5.1, the experimental group produced /p/ intelligibly in the pretest 28.7% of the time, while the control group did so 35.8% of the time. Thus, the groups showed similar results on the pretest. On the immediate posttest, the percentage of correct responses increased to 74.1% for the experimental group and only 59.4% for the control group. As a result, the experimental group showed a much bigger improvement in pronouncing /p/ than the control group. This could likely be attributed to the explicit pronunciation treatment.

On the delayed posttest, correct responses averaged 30.3% for the experimental group and 39.7% for the control group. This revealed that neither group maintained their improvement in pronouncing /p/ one month after the treatment. Figure 5.1 shows the likelihood of both groups giving a correct response in the pretest and immediate and delayed posttests.

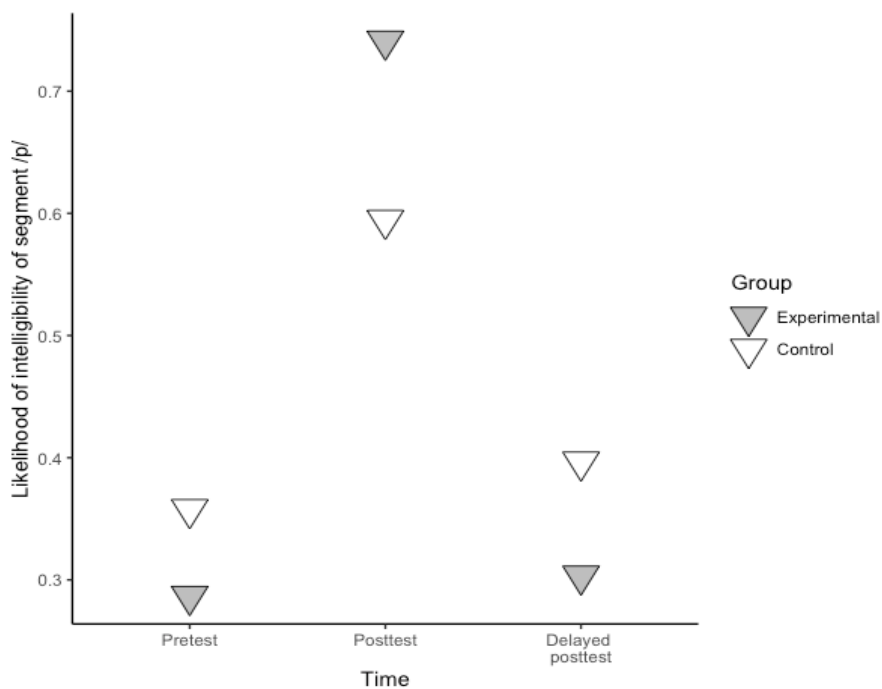


Figure 5.1. Likelihood of /p/ intelligibility.

### 5.2.1.2 Modelling Results for /p/ Intelligibility

To determine whether the observed difference after the treatment was significant, a generalised linear mixed effects model was fit in R (R Core Team, 2019) using lme4 (Bates et

al., 2015) and lmerTest (Kuznetsova et al., 2017). The perceived intelligibility of a correct response was added as a dependent variable. The fixed effects in the full model comprised a three-way interaction between group (experimental and control), time (pretest and immediate and delayed posttests) and the position of /p/ in each stimulus item (onset and coda). These factors served as predictors of response correctness, while participant, rater, and word were random effects. The formula for the full model for the intelligibility of /p/ was  $\text{Correctness} \sim \text{Group} * \text{Time} * \text{Position} + (1|\text{Word}) + (1|\text{Participant}) + (1|\text{Rater})$ . If an interaction or fixed effect was non-significant, the model was pruned until only significant predictors remained ( $p < 0.05$ ).

After pruning, the final model was  $\text{Correctness} \sim \text{Group} * \text{Time} + \text{Time} * \text{Position} + (1|\text{Word}) + (1|\text{Participant}) + (1|\text{Rater})$ . Table 5.2 represents the final model. The control group, pretest phase, and onset position were chosen as the reference levels. Time, group, and the two-way interactions between group and time and between time and position were significant predictors of the correctness of /p/.

Table 5.1: Model for /p/ Intelligibility with Pretest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.53585	0.09254	-5.790	7.02e-09	***
Group=Experimental	-0.32763	0.10150	-3.228	0.00125	**
Time=Posttest	0.86304	0.12039	7.169	7.58e-13	***
Time=Delayed.posttest	0.32416	0.11976	2.707	0.00680	**
Position=Coda	-0.10258	0.10134	-1.012	0.31142	
Group=Experimental Time=Posttest	0.99928	0.14401	6.939	3.95e-12	***
Group=Experimental Time=Delayed.posttest	-0.08983	0.14259	-0.630	0.52873	
Time=Posttest Position=Coda	0.21582	0.14343	1.505	0.13242	
Time=Delayed.posttest Position=Coda	-0.32473	0.14249	-2.279	0.02267	*

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

Raters were significantly more likely to judge participants in the control group as intelligible than the experimental group for the baseline. The control group in the immediate and delayed posttests was significantly more likely to be judged intelligible than in the pretest.

There was no effect of position for the baseline. The experimental group improved more than the control group from the pretest to the immediate posttest; however, the improvement from the pretest to the delayed posttest was similar for the two groups. A significant difference was also found in the two-way interaction between time and position, such that improvement in intelligibility from pretest to delayed posttest was less for coda position.

When the model was re-run with the posttest phase as the reference level, significant differences were found based on control group and onset position. In the immediate posttest, the experimental group was significantly more likely to be rated as intelligible than the control group. The control group was judged as less intelligible in the pretest and delayed posttest than in the immediate posttest. With respect to two-way interactions, the difference in intelligibility between immediate posttest and the other two tests was greater for the experimental group than the control group. Furthermore, regression in intelligibility from immediate to delayed posttest was significantly greater for the coda position than onset, as displayed in Table 5.3.

Table 5.3: Model for /p/ Intelligibility with Posttest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	0.32717	0.09147	3.577	0.000348	***
Group=Experimental	0.67165	0.10216	6.574	4.89e-11	***
Time=Pretest	-0.86299	0.12040	-7.168	7.64e-13	***
Time=Delayed.posttest	-0.53888	0.11895	-4.530	5.89e-06	***
Position=Coda	0.11326	0.10152	1.116	0.264534	
Group=Experimental Time=Pretest	-0.99928	0.14402	-6.939	3.96e-12	***
Group=Experimental Time=Delayed.posttest	-1.08912	0.14308	-7.612	2.70e-14	***
Time=Pretest Position=Coda	-0.21595	0.14344	-1.505	0.132202	
Time=Delayed.posttest Position=Coda	-0.54056	0.14262	-3.790	0.000151	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The final model for the segment /p/ is visualised in Figure 5.2.

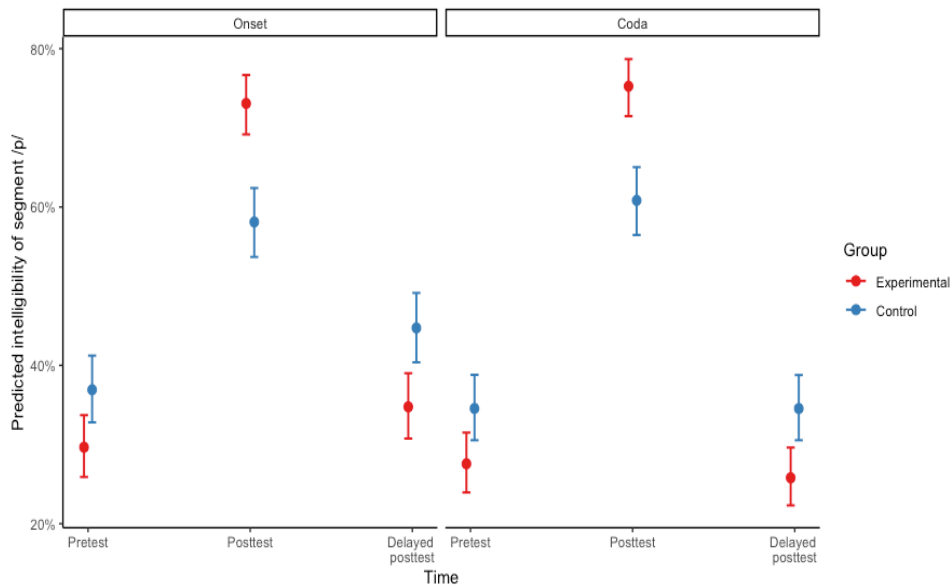


Figure 5.2. Model prediction of /p/ intelligibility based on time (pretest, posttests), group (experimental, control), and position (onset, coda).

## 5.2.2 The Segment /v/

### 5.2.2.1 Descriptive Results for /v/ Intelligibility

The conditions for /p/ were the same for /v/ in terms of occurrences containing the segment and how they were calculated. Table 5.4 presents the descriptive results for /v/ intelligibility.

Table 5.4: Descriptive Statistical Summary for /v/ Intelligibility

Group	Test Response					
	Pretest		Immediate Posttest		Delayed Posttest	
	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
Experimental	598 (66.4%)	302 (33.6%)	184 (20.4%)	716 (79.6%)	505 (58.4%)	395 (43.9%)
Control	526 (58.4%)	374 (41.6%)	496 (55.1%)	404 (44.9%)	569 (63.2%)	331 (36.8%)

In the pretest, the experimental group produced /v/ correctly 33.6% of the time, noticeably lower than the control group (41.6%). In the immediate posttest, correct responses by the experimental group showed a much higher increase to 79.6%, in contrast to the control

group, which went up slightly to 44.9%. Taken together, this improvement could be attributed to the explicit pronunciation treatment.

In the delayed posttest phase, correct responses fell to 43.9% for the experimental group and 36.8% for the control group. Thus, the gains made in pronouncing /v/ for both groups were not well-retained one month after the treatment, although the control group still showed lower overall attainment than the experimental group. Figure 5.3 demonstrates the likelihood of intelligibility in the pretest and posttests.

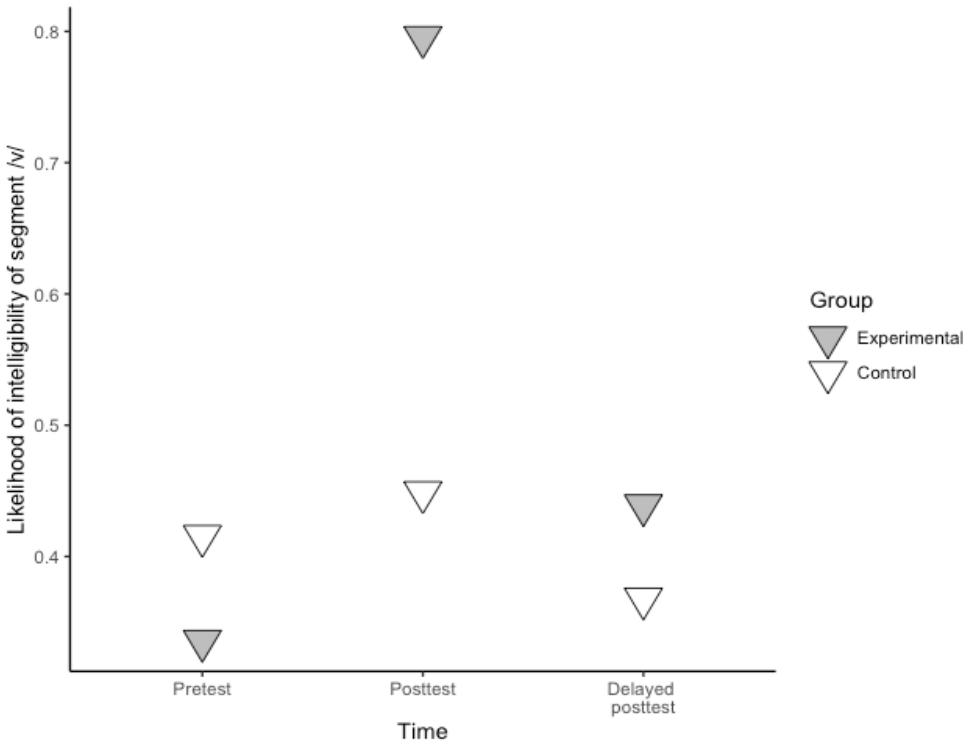


Figure 5.3. Likelihood of /v/ intelligibility.

### 5.2.2.2 Modelling Results for /v/ Intelligibility

A generalised linear mixed effects model was fit in R using the same formula and modelling procedures as those for /p/. After pruning, the final model for /v/ was  $\text{Correctness} \sim \text{Group} * \text{Time} + (1 | \text{Word}) + (1 | \text{Participant}) + (1 | \text{Rater})$ . Table 5.5 presents the final model. Unlike /p/, position was not a significant predictor with /v/. Therefore, it was excluded from the final

model. On the other hand, time, group, and the two-way interaction between group and time were significant predictors of intelligibility.

The control group was significantly more likely to be intelligible than the experimental group for the baseline. The control group in the immediate posttest was not significantly more intelligible than in the pretest but was more intelligible than in the pretest and delayed posttest. There was also a significant two-way interaction between group and time. The experimental group in the immediate and delayed posttests showed more improvement than the control group.

Table 5.5: Model for /v/ Intelligibility with Pretest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.34110	0.06803	-5.014	5.34e-07	***
Group=Experimental	-0.34216	0.09777	-3.499	0.000466	***
Time=Posttest	0.13592	0.09522	1.427	0.153470	
Time=Delayed.posttest	-0.20072	0.09672	-2.075	0.037958	*
Group=Experimental Time=Posttest	1.90623	0.14452	13.190	< 2e-16	***
Group=Experimental Time=Delayed.posttest	0.63828	0.13730	4.649	3.34e-06	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

As shown in Table 5.6, with the immediate posttest as the reference level, the experimental group was significantly more likely to be intelligible than the control group. The control group in the delayed posttest was significantly less intelligible than in the immediate posttest.

Table 5.6: Model for /v/ Intelligibility with Posttest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.20519	0.06742	-3.044	0.002338	**
Group=Experimental	1.56408	0.10642	14.697	< 2e-16	***
Time=Pretest	-0.13590	0.09522	-1.427	0.153541	
Time=Delayed.posttest	-0.33665	0.09629	-3.496	0.000472	***
Group=Experimental Time=Pretest	-1.90625	0.14452	-13.190	< 2e-16	***
Group=Experimental Time=Delayed.posttest	-1.26796	0.14359	-8.831	< 2e-16	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

With respect to the two-way interaction between time and group, the experimental group showed a larger drop in intelligibility in the pretest and delayed posttest compared to the control group. The interaction between time, group, and position is visualised in Figure 5.4 to predict /v/ intelligibility.

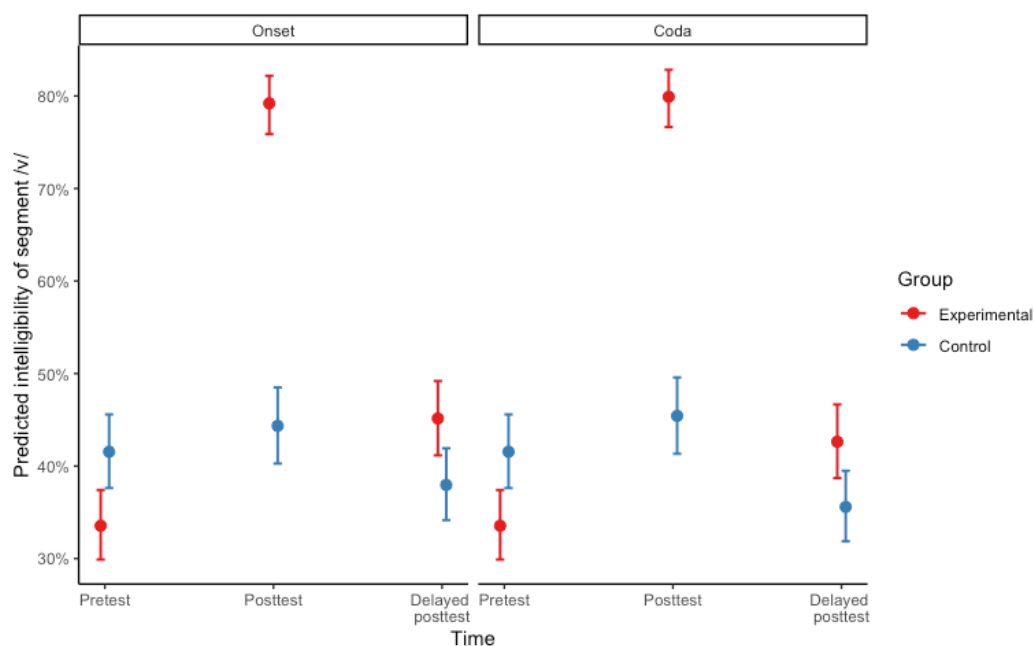


Figure 5.4. Model prediction of /v/ intelligibility based on time (pretest, posttests) and group (experimental, control).

### 5.2.3. The Segment /tʃ/

#### 5.2.3.1 Descriptive Results for /tʃ/ Intelligibility

The conditions for /tʃ/ were the same as previous segments in terms of total occurrences and how they were calculated (see Table 5.7).

Table 5.7: Descriptive Statistical Summary for /tʃ/ Intelligibility

Group	Test Response					
	Pretest		Immediate Posttest		Delayed Posttest	
	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
Experimental	664 (73.8%)	236 (26.2%)	507 (56.3%)	393 (43.7%)	575 (63.9%)	325 (36.1%)
Control	596 (66.2%)	304 (33.8%)	518 (57.6%)	382 (42.4%)	497 (55.2%)	403 (44.8%)

As shown in Table 5.7, the experimental group's rate of pronouncing /tʃ/ intelligibly in the pretest was 26.2%, while the control group was 33.8%. In the immediate posttest, this rate increased to 43.7% for the experimental group and 42.4% for the control group. That result revealed no sizeable improvement in the experimental group over the control group immediately after the explicit pronunciation treatment. In the delayed posttest, correct responses from the experimental group went down to 36.1%, while the control group went slightly up to 44.8%. The experimental group scored significantly lower than the control group, retaining little improvement they had achieved in the immediate posttest. Therefore, the explicit treatment played no role in pronouncing /tʃ/ intelligibly over time. Figure 5.5 displays the likelihood of intelligibility in the pretest and posttests.

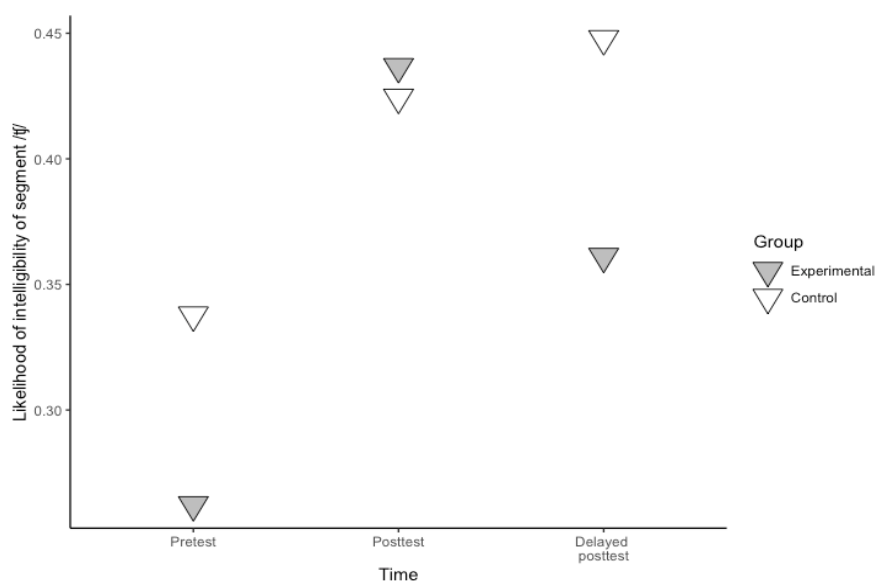


Figure 5.5. Likelihood of /tʃ/ intelligibility in pretest and posttests.

### 5.2.3.2 Modelling Results for /tʃ/ Intelligibility

A generalised linear mixed effects model was fit in R using the same formula as the full models for /p/ and /v/ and following the same modelling procedures. After pruning, the output of the final model was like that of /v/ in terms of significant predictors (see Table 5.8). The control group was judged as significantly more intelligible than the experimental group for the

baseline. The control group in the immediate and delayed posttests was significantly more likely to be intelligible than in the pretest. The experimental group improved significantly more in the immediate posttest than the control group. However, no significant difference was found in their improvement from the pretest to the delayed posttest.

Table 5.8: Model for /ʃ/ Intelligibility with Pretest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.690132	0.132747	-5.199	2.01e-07	***
Group=Experimental	-0.396919	0.171542	-2.314	0.020677	*
Time=Posttest	0.378072	0.098760	3.828	0.000129	***
Time=Delayed.posttest	0.475487	0.098496	4.827	1.38e-06	***
Group=Experimental Time=Posttest	0.438779	0.143201	3.064	0.002183	**
Group=Experimental Time=Delayed.posttest	0.009352	0.143955	0.065	0.948201	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

When the model was re-run with the immediate posttest as the reference level, the experimental group was not significantly different from the control group for the baseline. The control group in the delayed posttest was not significantly different from the immediate posttest. With respect to the two-way interaction between group and time, the experimental group showed a larger difference between the immediate posttest and the other two time points compared to the control group, as shown in Table 5.9.

Table 5.9: Model for /ʃ/ Intelligibility with Posttest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.31207	0.13112	-2.380	0.017314	*
Group=Experimental	0.04185	0.16650	0.251	0.801538	
Time=Pretest	-0.37807	0.09876	-3.828	0.000129	***
Time=Delayed.posttest	0.09740	0.09631	1.011	0.311901	
Group=Experimental Time=Pretest	-0.43879	0.14320	-3.064	0.002183	**
Group=Experimental Time=Delayed.posttest	-0.42940	0.13814	-3.108	0.001881	**

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between time, group, and position is plotted in Figure 5.6 to predict the degree of intelligibility of the segment /ʃ/.

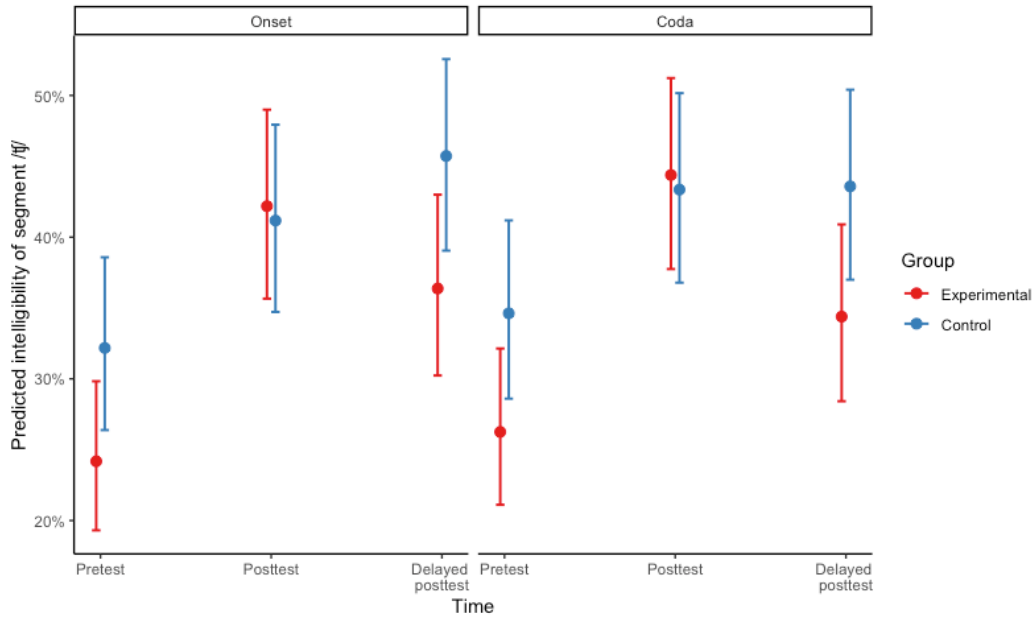


Figure 5.6. Model prediction of /ŋ/ intelligibility based on time (pretest, posttests) and group (experimental, control).

## 5.2.4 The Segment /ŋ/

### 5.2.4.1 Descriptive Results for /ŋ/ Intelligibility

The conditions for /ŋ/ were the same as the other segments in terms of occurrences including /ŋ/ and how they were calculated (see Table 5.10).

Table 5.10: Descriptive Statistical Summary for /ŋ/ Intelligibility

Group	Test Response					
	Pretest		Immediate Posttest		Delayed Posttest	
	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
Experimental	678 (75.3%)	222 (24.7%)	402 (44.7%)	498 (55.3%)	585 (65%)	315 (35%)
Control	617 (68.6%)	283 (31.4%)	640 (71.1%)	260 (28.9%)	546 (60.7%)	354 (39.3%)

In Table 5.10, the experimental group's mean intelligibility of /ŋ/ in the pretest was 24.7%, while it was higher for the control group (31.4%). The groups were somewhat similar prior to the treatment. In the immediate posttest, intelligible responses more than doubled to 55.3% for the experimental group but declined slightly for the control group to 28.9%. Overall,

this improvement could be attributed to the explicit treatment. In the delayed posttest, mean intelligibility was 35% for the experimental group and 39.3% for the control group. Thus, the experimental group did not retain its gains with /ŋ/ one month after the treatment. In contrast, the control group showed some improvement in the immediate and delayed posttests (see Figure 5.7).

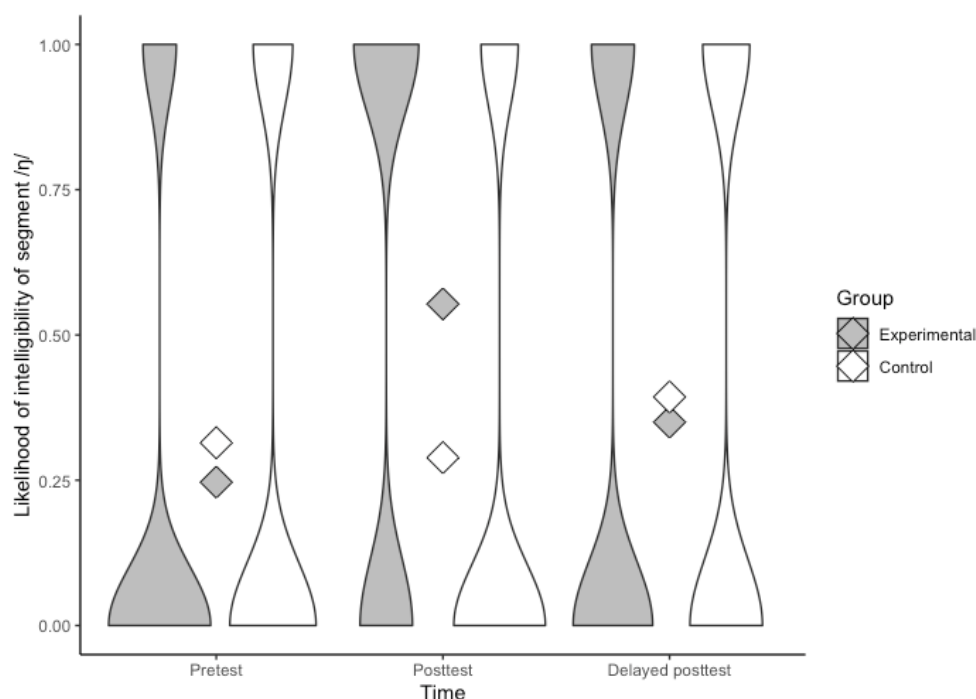


Figure 5.7. Likelihood of /ŋ/ intelligibility in pretest and posttests.

#### 5.2.4.2 Modelling Results for /ŋ/ Intelligibility

A generalised linear mixed effects model was fit in R by applying the same formula as the full models of the other segments with the same modelling procedures, except that the predictor of position was not included in the full model. After pruning, the final model was similar to /v/ and /ʃ/ in terms of the significant predictors being time, group, and the two-way interaction between time and group (see Table 5.11).

Table 5.11: Model for /ŋ/ Intelligibility with Pretest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.8005	0.1124	-7.121	1.07e-12	***
Group=Experimental	-0.3339	0.1485	-2.249	0.024537	*
Time=Posttest	-0.1246	0.1041	-1.198	0.231051	
Time=Delayed.posttest	0.3560	0.1004	3.545	0.000392	***
Group=Experimental Time=Posttest	1.4772	0.1466	10.075	< 2e-16	***
Group=Experimental Time=Delayed.posttest	0.1489	0.1452	1.025	0.305228	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The control group and pretest served as the reference level and were found to be significant predictors of the intelligibility of /ŋ/. The experimental group was significantly less likely to be intelligible than the control group in the pretest, while the control group in the immediate posttest was not significantly more likely to be more intelligible than in the pretest. However, the control group was significantly more likely to attain more intelligible speech in the delayed posttest than in the pretest. A significant difference was found in the two-way interaction between group and time. The experimental group in the immediate posttest was significantly more likely to achieve intelligible speech than the control group. Nevertheless, the experimental group failed to achieve significant improvement over the control group in the delayed posttest.

The control group and immediate posttest as the reference level also revealed a significant difference. The experimental group in the immediate posttest was significantly more likely to be intelligible than the control group, while the control group's performance in the delayed posttest was significantly different from the immediate posttest. With respect to the two-way interaction between time and group, the experimental group showed a significant difference in the pretest and regression in the delayed posttest compared to the control group (see Table 5.12).

Table 5.12: Model for /ŋ/ Intelligibility with Posttest as Reference Level

	Estimate	Std. Error	z-Value	Pr(> z )	Sig.
(Intercept)	-0.9252	0.1136	-8.145	3.79e-16	***
Group=Experimental	1.1434	0.1443	7.922	2.34e-15	***
Time=Pretest	0.1246	0.1041	1.198	0.231	
Time=Delayed.posttest	0.4806	0.1017	4.726	2.29e-06	***
Group=Experimental Time=Pretest	-1.4772	0.1466	-10.075	< 2e-16	***
Group=Experimental Time=Delayed.posttest	-1.3283	0.1411	-9.416	< 2e-16	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The final model for /ŋ/ is illustrated in Figure 5.8.

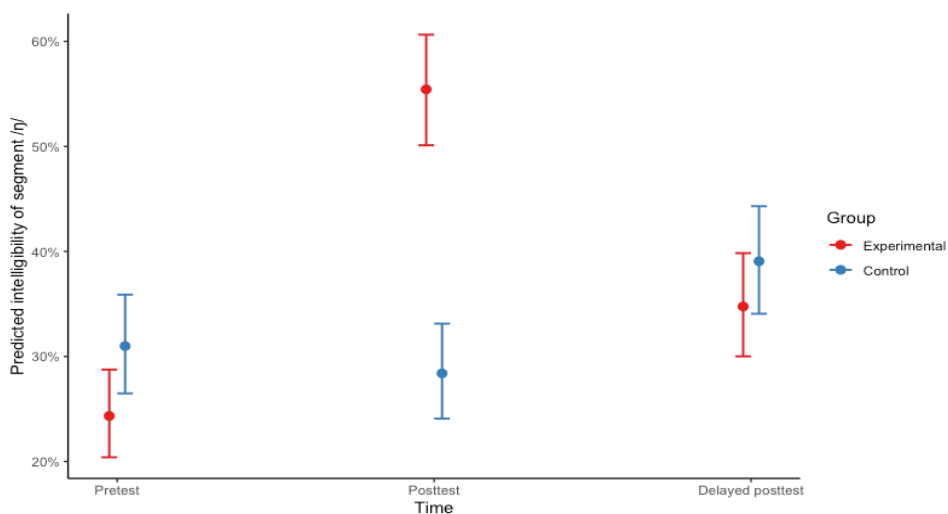


Figure 5.8. Model prediction for /ŋ/ intelligibility based on time (pretest, posttests) and group (experimental, control).

### 5.3 Comprehensibility of Segmentals

This section reports the descriptive and modelling results regarding the comprehensibility of segmental features.

#### 5.3.1 Descriptive Results

The study compared the experimental and control groups' mean scores in the pretest and immediate and delayed posttests, as shown in Table 5.13.

Table 5.13: Descriptive Statistical Summary for Comprehensibility of Segmentals

Group	Test					
	Pretest		Immediate Posttest		Delayed Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental	3.16	0.98	5.41	1.08	4.38	0.94
Control	2.88	1.15	4.6	1.29	3.90	1.35

Although the mean score of the experimental group in the pretest ( $M = 3.16$ ) was higher than the control group ( $M = 2.88$ ), no substantial differences between groups in the pretest were found. This indicated that any differences between scores in the immediate and delayed posttests could be attributed to the intervention. In the immediate posttest, the experimental group’s mean score ( $M = 5.41$ ) was higher than the control group ( $M = 4.6$ ), suggesting the effectiveness of the intervention. In the delayed posttest, the experimental group still achieved a higher mean ( $M = 4.38$ ) than the pretest but lower than the immediate posttest. The mean comprehensibility of the control group was ( $M = 3.90$ ), which was also less than the immediate posttest. This means both groups failed to retain gains in comprehensibility over time (see Figure 5.9).

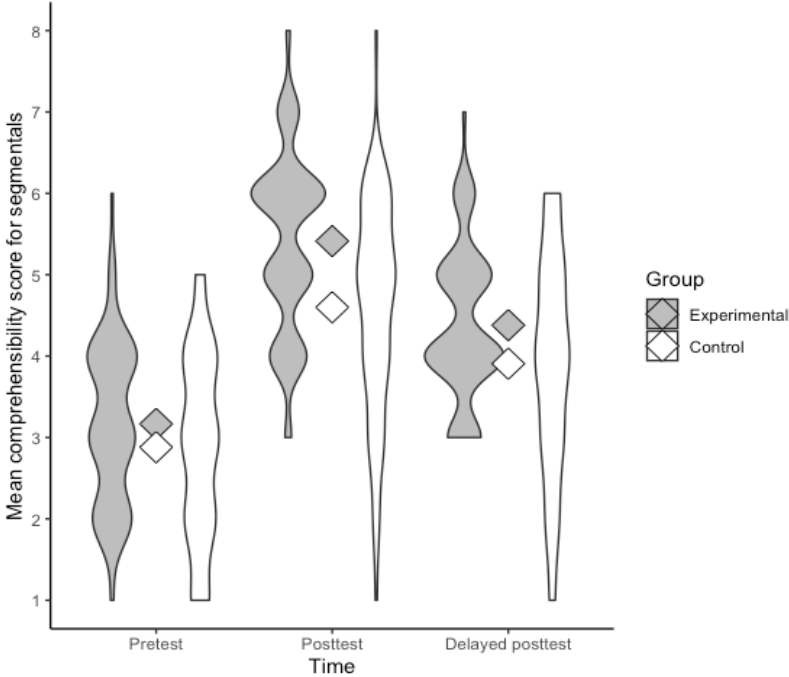


Figure 5.9. Mean comprehensibility scores for segmentals in pretest and posttests.

### 5.3.2 Modelling Results for Comprehensibility of Segmentals

A linear mixed effects model was fit in R. The comprehensibility score served as a dependent variable, and the fixed effects in the full model included a two-way interaction between group (experimental, control) and time (pretest, posttests) as predictors. Participant and rater were random effects. The formula for the full model for perceived comprehensibility was  $\text{Score} \sim \text{Group} * \text{Time} + (1|\text{Participant}) + (1|\text{Rater})$ . Table 5.14 presents the final model.

The control group and pretest were assigned as the reference level. For the intercept (control group and pretest), the predicted comprehensibility score was 2.8800. Time, group, and the two-way interaction between time and group were significant predictors of comprehensibility. The experimental group was not significantly different from the control group in the pretest, meaning any difference in the posttests could be attributed to the intervention. The control group was significantly more likely to be comprehensible in the immediate and delayed posttests than in the pretest. Based on the interaction between group and time, the experimental group was significantly more likely to be comprehensible in the immediate posttest but not in the delayed posttest.

Table 5.14: Model for Comprehensibility of Segmentals with Pretest as Reference Level

	Estimate	Std. Error	Df	t-Value	Pr(> t )	Sig.
(Intercept)	2.8800	0.1676	39.6726	17.183	< 2e-16	***
Group=Experimental	0.2867	0.2370	39.6726	1.209	0.23366	
Time=Posttest	1.7200	0.1163	866.0000	14.794	< 2e-16	***
Time=Delayed.posttest	1.0267	0.1163	866.0000	8.831	< 2e-16	***
Group= Experimental Time=Posttest	0.5267	0.1644	866.0000	3.203	0.00141	**
Group=Experimental Time=Delayed.posttest	0.1867	0.1644	866.0000	1.135	0.25656	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

With the control group and immediate posttest as the reference level, significant differences were found again (see Table 5.15). The experimental group was more comprehensible than the control group in the immediate posttest, and the control group

demonstrated lower comprehensibility in both the pretest and delayed posttest. This difference was even larger for the experimental group with respect to the interaction between group and time.

Table 5.15: Model for Comprehensibility of Segmentals with Posttest as Reference Level

	Estimate	Std. Error	Df	t-Value	Pr(> t )	Sig.
(Intercept)	4.6000	0.1676	39.6726	27.446	< 2e-16	***
Group=Experimental	0.8133	0.2370	39.6726	3.431	0.00142	**
Time=Pretest	-1.7200	0.1163	866.0000	-14.794	< 2e-16	***
Time=Delayed.posttest	-0.6933	0.1163	866.0000	-5.964	3.59e-09	***
Group= Experimental Time=Pretest	-0.5267	0.1644	866.0000	-3.203	0.00141	**
Group=Experimental Time=Delayed.posttest	-0.3400	0.1644	866.0000	-2.068	0.03895	*

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between the experimental and control groups and the pretest and posttests is illustrated in Figure 5.10.

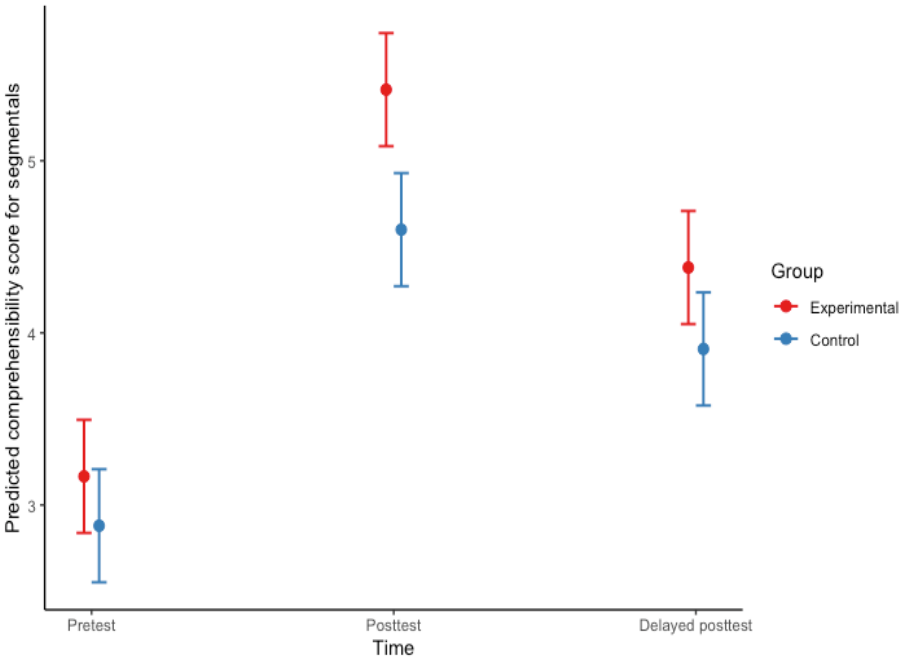


Figure 5.10. Model prediction for the comprehensibility of segmentals based on time (pretest, posttests) and group (experimental, control).

**5.4 Intelligibility of Suprasegmentals**

This section reports the suprasegmental results regarding intelligibility, including lexical stress, rhythm, and final intonation.

**5.4.1 Lexical Stress**

**5.4.1.1 Descriptive Results for Intelligibility of Lexical Stress**

Ten stimulus items targeting lexical stress were recorded by 30 participants in the experimental and control groups in the pretest and posttests. These recordings were assessed by 10 raters, yielding 9,000 occurrences (30 participants × 3 tests × 10 words × 10 raters). Table 5.16 presents a summary of the descriptive results.

Table 5.16: Descriptive Statistical Summary for Intelligibility of Lexical Stress

Group	Test Response					
	Pretest		Immediate Posttest		Delayed Posttest	
	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
Experimental	1,133 (75.5%)	367 (24.5%)	494 (32.9%)	1,006 (67.1%)	541 (36.1%)	959 (63.9%)
Control	1,050 (70%)	450 (30%)	888 (59.2%)	612 (40.8%)	997 (66.5%)	503 (33.5%)

In the pretest, 24.5% of the experimental group’s responses were judged to be intelligible, lower than the control group (30%). The similar rates of correct responses in the pretest indicated that any differences in the immediate and delayed posttests could be attributed to the intervention. In the immediate posttest, correct responses rose to 67.1% for the experimental group and only 40.8% for the control group. This remarkable improvement over the control group could likely be attributed to the explicit pronunciation intervention. In the delayed posttest, correct responses fell only slightly to 63.9% for the experimental group and 33.5% for the control group. The experimental group thus largely retained pronunciation gains in lexical stress one month after the explicit intervention. Figure 5.11 illustrates the likelihood of correctness in the pretest and posttests.

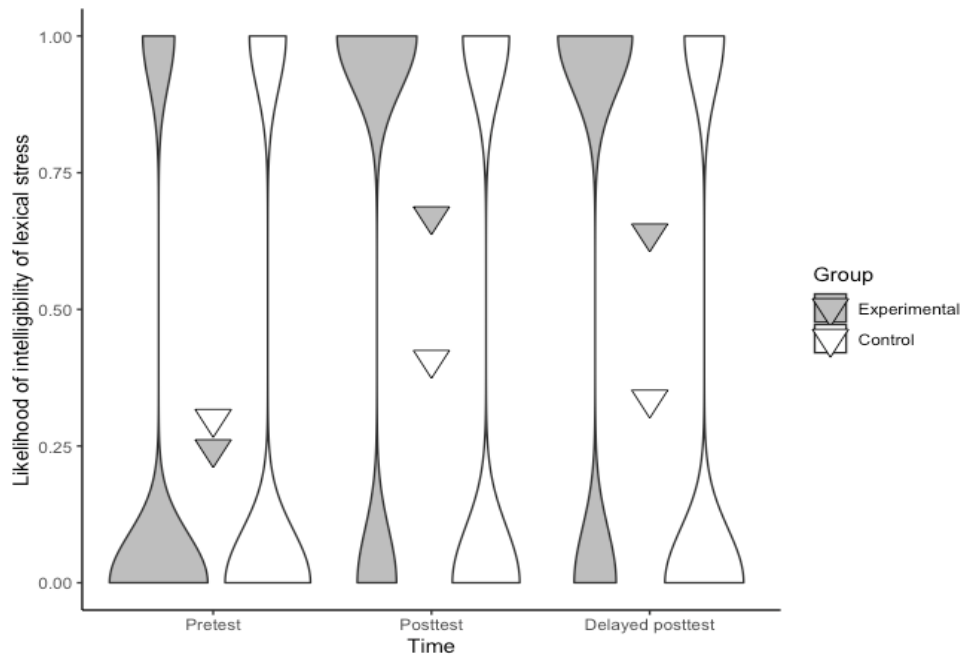


Figure 5.11. Likelihood of intelligibility of lexical stress.

#### 5.4.1.2 Modelling Results for Intelligibility of Lexical Stress

A generalised linear mixed effects model was fit in R using the same formula as the segmental full models and following the same modelling procedures but without position as a fixed effect in the full model. Time, group, and the two-way interaction between time and group were significant predictors of lexical stress intelligibility. Table 5.17 presents the final model.

Table 5.17: Model for Intelligibility of Lexical Stress with Pretest as Reference Level

	Estimate	Std. Error	<i>z</i>	Pr(>  <i>z</i>  )	Sig.
(Intercept)	-0.86386	0.08527	-10.131	< 2e-16	***
Group=Experimental	-0.27118	0.11956	-2.268	0.0233	*
Time=Posttest	0.48466	0.07781	6.229	4.71e-10	***
Time=Delayed.posttest	0.16633	0.07926	2.099	0.0359	*
Group=Experimental Time=Posttest	1.36677	0.11284	12.112	< 2e-16	***
Group=Experimental Time=Delayed.posttest	1.54541	0.11329	13.641	< 2e-16	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The control group was significantly more intelligible than the experimental group for the baseline. The control group was also significantly more likely to be more intelligible in the immediate and delayed posttests than in the pretest. The experimental group in the immediate

and delayed posttests was also significantly more likely to attain more intelligible speech. This might be attributed to the pronunciation intervention.

With the immediate posttest as the reference level, a significant difference between the two groups was found. Additionally, a significant negative difference was found with the pretest and delayed posttest for the control group. The experimental group also showed a comparatively larger difference with the pretest (see Table 5.18).

Table 5.18: Model for Intelligibility of Lexical Stress with Posttest as Reference Level

	Estimate	Std. Error	z	Pr(> z )	Sig.
(Intercept)	-0.37919	0.08272	-4.584	4.57e-06	***
Group=Experimental	1.09558	0.11522	9.509	< 2e-16	***
Time=Pretest	-0.48467	0.07780	-6.229	4.68e-10	***
Time=Delayed.posttest	-0.31833	0.07659	-4.156	3.24e-05	***
Group=Experimental Time=Pretest	-1.36675	0.11282	-12.114	< 2e-16	***
Group=Experimental Time=Delayed.posttest	0.17865	0.10869	1.644	0.1	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between time and group is plotted in Figure 5.12.

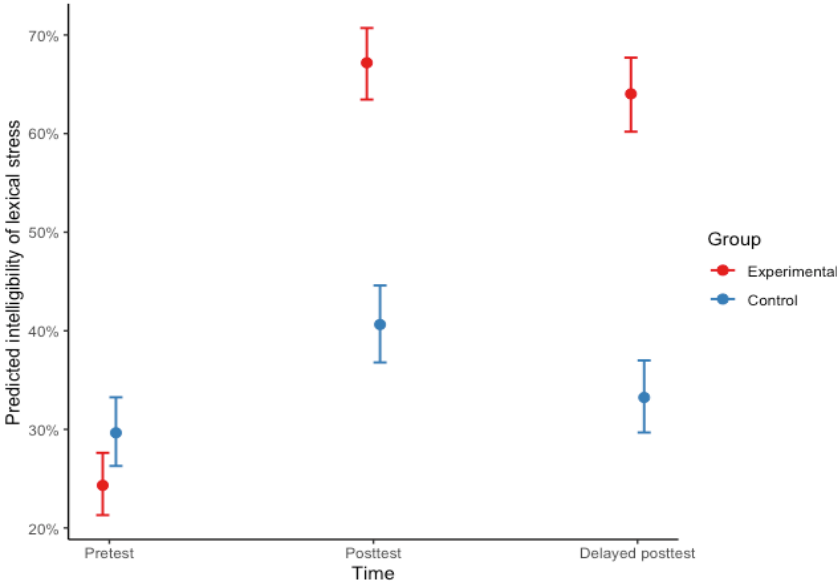


Figure 5.12. Model prediction for intelligibility of lexical stress based on time (pretest, posttests) and group (experimental, control).

## 5.4.2 Rhythm

### 5.4.2.1 Descriptive Results for Intelligibility of Rhythm

Mean scores for rhythm intelligibility from the pretest and posttests are compared in Table 5.19.

Table 5.19: Descriptive Statistical Summary for Intelligibility of Rhythm

Group	Test					
	Pretest		Immediate Posttest		Delayed Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental	0.17	0.23	0.18	0.24	0.16	0.22
Control	0.13	0.22	0.13	0.21	0.12	0.21

The experimental group had nearly the same mean intelligibility score in the pretest ( $M = 0.17$ ) as the control group ( $M = 0.13$ ). Therefore, any observed significant differences between the groups' scores in the immediate and delayed posttests could be attributed to the intervention. The experimental group's mean score in the immediate posttest ( $M = 0.18$ ) was almost the same as the pretest, while the control group had the same mean score as before ( $M = 0.13$ ). This finding suggested the intervention was ineffective at improving rhythm. In the delayed posttest, the mean intelligibility of the experimental group ( $M = 0.16$ ) and control group ( $M = 0.12$ ) slightly declined, with no apparent improvement a month after the treatment (see Figure 5.13).

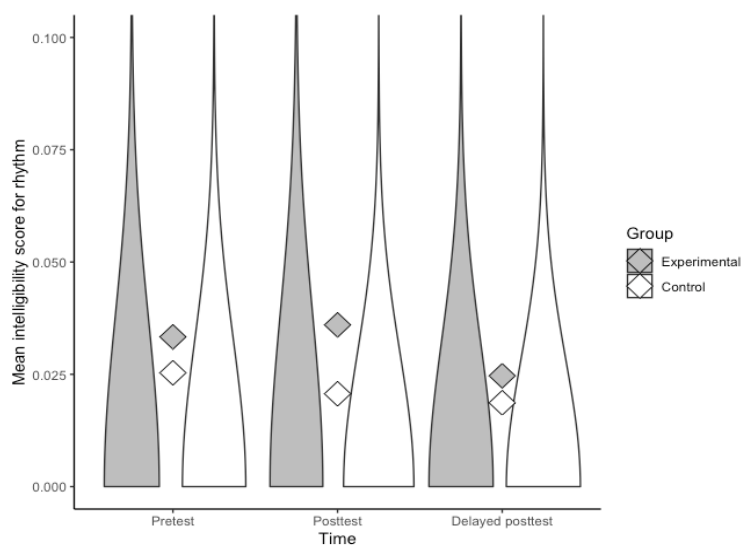


Figure 5.13. Mean intelligibility scores for rhythm.

### 5.4.2.2 Modelling Results for Intelligibility of Rhythm

A linear mixed effects model was fit in R with intelligibility as a dependent variable. The fixed effects in the full model included a three-way interaction between group (experimental and control), time (pretest and posttests), and sentence type (A and B). Participant, sentence, target word, and rater were random effects. The formula for the full model for rhythm intelligibility was  $\text{Score} \sim \text{Group} * \text{Time} * \text{Sentence.type} + (1|\text{Target.word}) + (1|\text{Sentence}) + (1|\text{Participant}) + (1|\text{Rater})$ . After pruning, the final model was  $\text{Score} \sim \text{Group} * \text{Time} + (1|\text{Target.word}) + (1|\text{Sentence}) + (1|\text{Participant}) + (1|\text{Rater})$ . For the intercept (control group and pretest), the predicted intelligibility score was 0.03195. Time, group, and a two-way interaction between them were not significant predictors of rhythm intelligibility. Table 5.20 presents the output of the final model.

Table 5.20: Model for Intelligibility of Rhythm with Pretest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	0.03195	0.03329	3.087	0.960	0.406	
Group=Experimental	0.008	0.005588	228.9	1.432	0.154	
Time=Posttest	-0.00667	0.005526	8954	-0.844	0.398	
Time=Delayed.posttest	-0.006667	0.005526	8954	-1.206	0.228	
Group=Experimental Time=Posttest	0.007333	0.007815	8954	0.938	0.348	
Group=Experimental Time=Delayed.posttest	-0.002	0.007815	8954	-0.256	0.798	

*Note.*  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The experimental group was not significantly different from the control group in the pretest, suggesting that any significant difference in the immediate and delayed posttests could be attributed to the treatment. The control group did not show significant differences in the immediate and delayed posttests compared to the pretest. The experimental group was also significantly less likely to achieve more intelligible speech in the posttests than in the pretest.

With the control group and immediate posttest as the reference level and a predicted intelligibility score of 0.02728, the experimental group was significantly more likely to be more intelligible than the control group in the immediate posttest, as illustrated in Table 5.21.

Table 5.21: Model for Intelligibility of Rhythm with Posttest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	0.02728	0.03329	3.086	0.819	0.47107	
Group=Experimental	0.01533	0.005588	228.8	2.744	0.00655	**
Time=Pretest	0.004667	0.005526	8954	0.844	0.39843	
Time=Delayed.posttest	-0.002	0.005526	8954	-0.362	0.71742	
Group=Experimental Time=Pretest	-0.007333	0.007815	8954	-0.938	0.34809	
Group=Experimental Time=Delayed.posttest	-0.009333	0.007815	8954	-1.194	0.23241	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between group and time is illustrated in Figure 5.14.

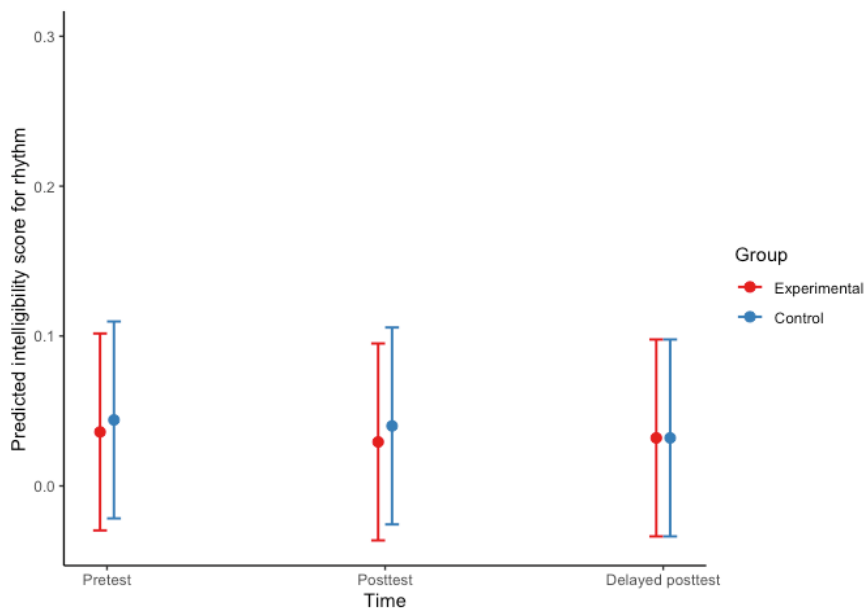


Figure 5.14. Model prediction for intelligibility of rhythm based on time (pretest, posttests) and group (experimental, control).

### 5.4.3 Final Intonation

#### 5.4.3.1 Descriptive Results for Intelligibility of Final Intonation

Instances of final intonation were calculated following the same methods as lexical stress (see Table 5.22).

Table 5.22: Descriptive Statistical Summary for Intelligibility of Final Intonation

Group	Test Response					
	Pretest		Immediate Posttest		Delayed Posttest	
	Incorrect	Correct	Incorrect	Correct	Incorrect	Correct
Experimental	1,065 (71%)	435 (29%)	447 (29.8%)	1,053 (70.2%)	480 (32%)	1,020 (68%)
Control	1,098 (73.2%)	402 (26.8%)	780 (52%)	720 (48%)	987 (65.8%)	513 (34.2%)

In the pretest, the experimental group gave correct responses 29% of the time, and the control group had a slightly lower rate of 26.8%. This similarity meant differences in the posttests might be a result of the pronunciation intervention (see Figure 5.15).

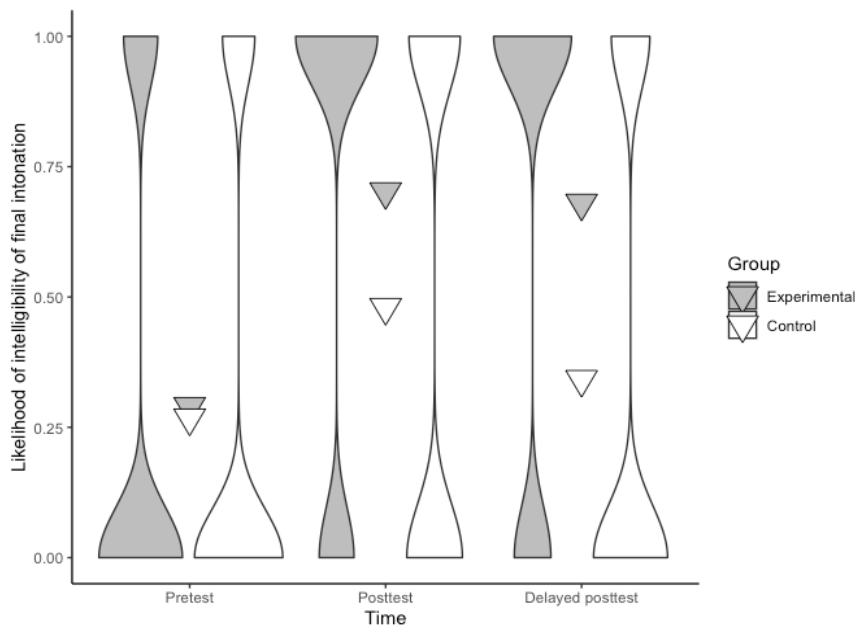


Figure 5.15. Likelihood of intelligibility of final intonation.

In the immediate posttest, the experimental group gave correct responses 70.2% of the time, and the control group gave them 48% of the time. This substantial improvement in the

intelligibility of final intonation likely resulted from the intervention. In the delayed posttest, correct responses reached 68% for the experimental group and 34.2% for the control group. Thus, the experimental group largely retained pronunciation improvements a month after the intervention.

### 5.4.3.2 Modelling Results for Intelligibility of Final Intonation

A generalised linear mixed effects model was fit in R using the same formula and modelling procedures as before but without position as a fixed effect and with the random effect “word” replaced with “sentence” in the full model. Therefore, the full model for final intonation was formulated as follows:  $\text{Correctness} \sim \text{Group} * \text{Time} + (1|\text{Participant}) + (1|\text{Sentence}) + (1|\text{Rater})$ . Time, group, and the two-way interaction between them were significant predictors of intelligible final intonation. Table 5.23 presents the output of the final model.

Table 5.23: Model for Intelligibility of Final Intonation with Pretest as Reference Level

	Estimate	Std. Error	z	Pr(> z )	Sig.
(Intercept)	-1.06554	0.15018	-7.095	1.29e-12	***
Group=Experimental	0.13768	0.13678	1.007	0.314	
Time=Posttest	0.98059	0.08035	12.204	< 2e-16	***
Time=Delayed.posttest	0.37006	0.08194	4.516	6.29e-06	***
Group=Experimental Time=Posttest	0.83472	0.11443	7.295	2.99e-13	***
Group=Experimental Time=Delayed.posttest	1.33874	0.11513	11.628	< 2e-16	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The experimental group was not significantly different from the control group for the reference. The control group in both posttests was significantly more likely to attain more intelligible speech than in the pretest. The experimental group was also more intelligible in the posttests than the pretest.

Relevelling was used to check for significant differences between groups when the immediate posttest was assigned as the reference level (see Table 5.24). In the immediate posttest, the experimental group was significantly more likely to be intelligible than the control

group. Intelligibility was lower in the pretest and delayed posttest for the control group. Additionally, it was comparatively lower in the pretest and higher in the delayed posttest for the experimental group.

Table 5.24: Model for Intelligibility of Final Intonation with Posttest as Reference Level

	Estimate	Std. Error	z	Pr(> z )	Sig.
(Intercept)	-0.08494	0.14759	-0.576	0.565	
Group=Experimental	0.97243	0.13385	7.265	3.73e-13	***
Time=Pretest	-0.98058	0.08035	-12.203	< 2e-16	***
Time=Delayed.posttest	-0.61052	0.07743	-7.885	3.16e-15	***
Group=Experimental Time=Pretest	-0.83474	0.11443	-7.295	2.99e-13	***
Group=Experimental Time=Delayed.posttest	0.50400	0.11158	4.517	6.27e-06	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between time and group in terms of final intonation is illustrated in Figure 5.16.

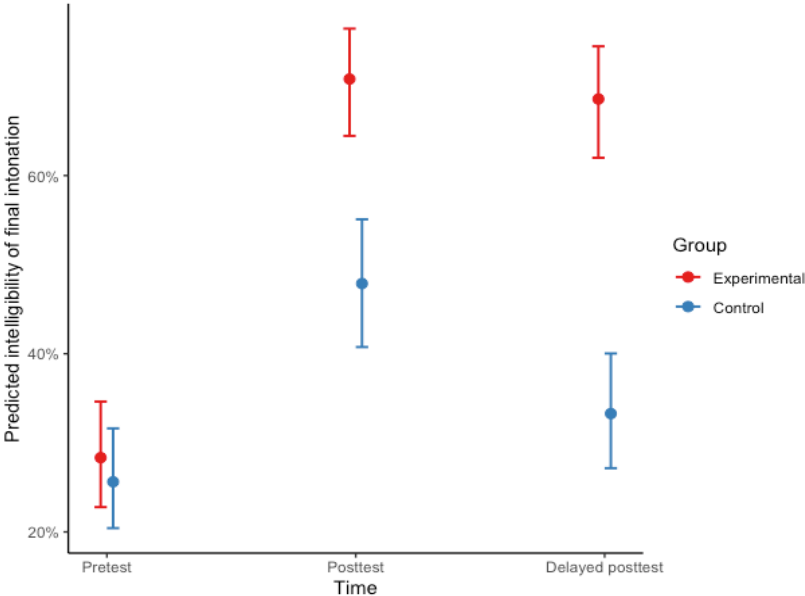


Figure 5.16. Model prediction for intelligibility of final intonation based on time (pretest, posttests) and group (experimental, control).

## 5.5 Comprehensibility of Suprasegmentals

This section reports the suprasegmental results for the comprehensibility of lexical stress, rhythm, and final intonation.

### 5.5.1 Lexical Stress

#### 5.5.1.1 Descriptive Results for Comprehensibility of Lexical Stress

The descriptive results for the comprehensibility of lexical stress are summarised in Table 5.25.

Table 5.25: Descriptive Statistical Summary for the Comprehensibility of Lexical Stress

Group	Test					
	Pretest		Immediate Posttest		Delayed Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental	2.60	0.88	7.13	0.85	6.16	0.89
Control	2.84	1.06	6.74	1.01	5.96	1.13

The experimental group's mean score in the pretest ( $M = 2.60$ ) was slightly lower than that of the control group ( $M = 2.84$ ). Figure 5.17 visually compares these scores.

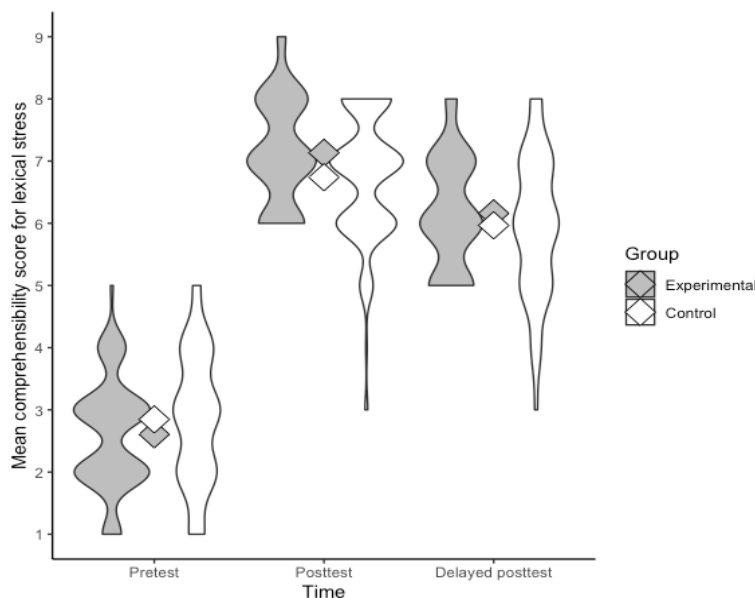


Figure 5.17. Mean comprehensibility scores of lexical stress.

In the immediate posttest, however, the experimental group scored a higher mean ( $M = 7.13$ ) than the control group ( $M = 6.74$ ). This could reflect the positive influence of the

explicit intervention. Although the experimental group’s mean score ( $M = 6.16$ ) went down in the delayed posttest, it was still higher than the control group ( $M = 5.96$ ).

**5.5.1.2 Modelling Results for Comprehensibility of Lexical Stress**

A linear mixed effects model was fit in R using the same formula as the full model for segmental comprehensibility. For the intercept (control group and pretest), the predicted comprehensibility score was 2.8467. As with segmental comprehensibility, time, group, and their two-way interaction were significant predictors of the comprehensibility of lexical stress. Table 5.26 shows the output of the final model. The experimental group was not significantly different from the control group in the pretest, meaning any difference in the posttests could be attributed to the intervention. The control group was significantly more likely to be more comprehensible in the posttests than in the pretest.

Table 5.26: Model for Comprehensibility of Lexical Stress with Pretest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(> t )	Sig.
(Intercept)	2.8467	0.1122	56.1865	25.368	< 2e-16	***
Group=Experimental	-0.2400	0.1585	58.0160	-1.514	0.13539	
Time=Posttest	3.8933	0.1076	856.9996	36.168	< 2e-16	***
Time=Delayed.posttest	3.1200	0.1076	856.9996	28.984	< 2e-16	***
Group=Experimental Time=Posttest	0.6333	0.1522	856.9996	4.160	3.5e-05	***
Group=Experimental Time=Delayed.posttest	0.4333	0.1522	856.9996	2.847	0.00453	**

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

With the immediate posttest as the reference level, the predicted comprehensibility score was 6.7400. The experimental group showed significantly higher comprehensibility in the immediate posttest. Comprehensibility in the pretest and delayed posttest was lower for the control group, while the difference between pretest and immediate posttest for the experimental group was still larger, as shown in Table 5.27.

Table 5.27: Model for Comprehensibility of Lexical Stress with Posttest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	6.7400	0.1122	56.1865	60.062	< 2e-16	***
Group=Experimental	0.3933	0.1585	58.0160	2.482	0.016	*
Time=Pretest	-3.8933	0.1076	856.9996	-36.168	< 2e-16	***
Time=Delayed.posttest	-0.7733	0.1076	856.9996	-7.184	1.47e-12	***
Group=Experimental Time=Pretest	-0.6333	0.1522	856.9996	-4.160	3.50e-05	***
Group=Experimental Time=Delayed.posttest	-0.2000	0.1522	856.9996	-1.314	0.189	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between group and time regarding lexical stress is visualised in Figure 5.18.

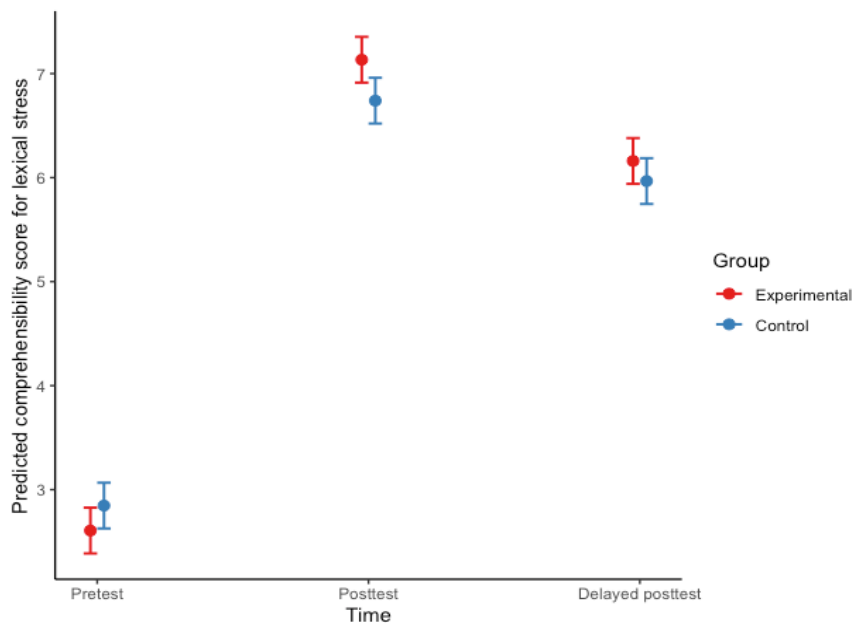


Figure 5.18. Model prediction for comprehensibility of lexical stress based on time (pretest, posttests) and group (experimental, control).

## 5.5.2 Rhythm

### 5.5.2.1 Descriptive Results for Comprehensibility of Rhythm

This section presents the descriptive results for rhythm comprehensibility. Table 5.28 shows the mean scores of the experimental and control groups in the pretest and posttests.

Table 5.28: Descriptive Statistical Summary for Comprehensibility of Rhythm

Group	Test					
	Pretest		Immediate Posttest		Delayed Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental	2.04	0.77	3.79	0.70	2.94	0.94
Control	2.08	0.78	3.08	0.72	2.32	0.95

The mean score of the experimental group in the pretest ( $M = 2.04$ ) was nearly the same as that of the control group ( $M = 2.08$ ). Thus, any difference in the posttests could likely be attributed to the intervention. In the immediate posttest, the mean score of the experimental group ( $M = 3.79$ ) was higher than the control group ( $M = 3.08$ ). This suggested the explicit intervention was effective. In the delayed posttest, the experimental group’s mean score dropped ( $M = 2.94$ ), as did that of the control group ( $M = 2.32$ ). Figure 5.19 compares these scores visually.

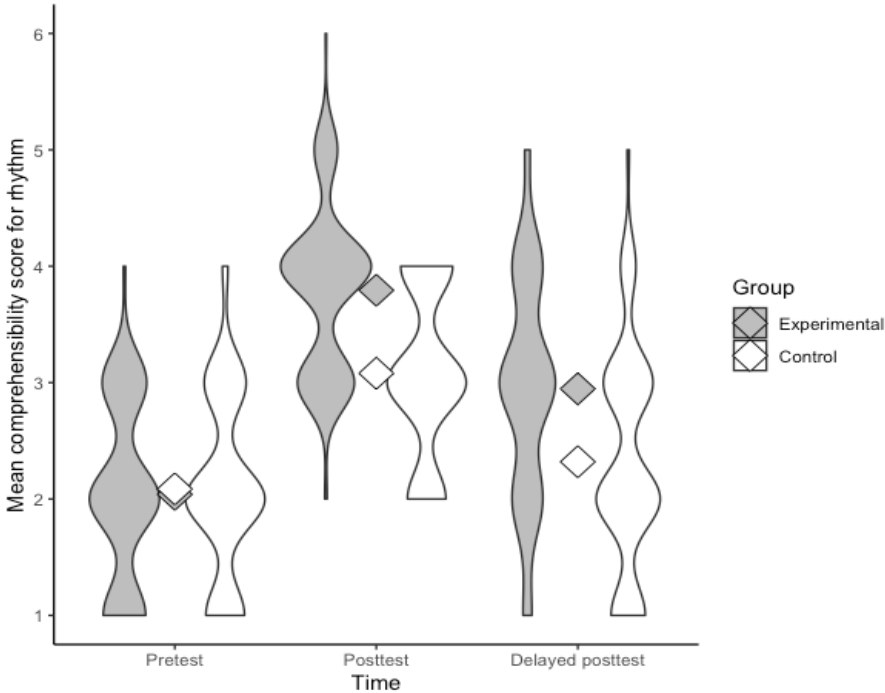


Figure 5.19. Mean comprehensibility scores for rhythm.

### 5.5.2.2 Modelling Results for Comprehensibility of Rhythm

A linear mixed effects model was fit in R using the same formula as the full model for segmental comprehensibility and lexical stress. For the intercept (control group and pretest), the predicted comprehensibility score was 2.08667. Like comprehensibility for segmentals and lexical stress, time, group, and their two-way interaction were found to be significant predictors of comprehensible rhythm. Table 5.29 presents the output of the final model.

The experimental group was not significantly different from the control group in the pretest, meaning any significant difference in the posttests could be attributed to the intervention. The control group was significantly more likely to achieve more comprehensible speech in the immediate and delayed posttests than in the pretest. This regression was even more apparent in the experimental group.

Table 5.29: Model for Comprehensibility of Rhythm with Pretest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	2.08667	0.10086	50.62342	20.689	< 2e-16	***
Group=Experimental	-0.04667	0.14171	50.93793	-0.329	0.74326	
Time=Posttest	0.99333	0.08851	857.00006	11.223	< 2e-16	***
Time=Delayed.posttest	0.23333	0.08851	857.00006	2.636	0.00853	**
Group=Experimental Time=Posttest	0.76000	0.12517	857.00006	6.072	1.90e-09	***
Group=Experimental Time=Delayed.posttest	0.67333	0.12517	857.00006	5.379	9.65e-08	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

With the control group and immediate posttest as the reference level, the predicted comprehensibility score was 3.08000. The experimental group demonstrated higher comprehensibility, while it was significantly lower in the pretest and delayed posttest for the control group. The difference with the pretest was even larger for the experimental group, as presented in Table 5.30.

Table 5.30: Model for Comprehensibility of Rhythm with Posttest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	3.08000	3.08000	50.62342	30.538	< 2e-16	***
Group=Experimental	0.71333	0.14171	50.93793	5.034	6.37e-06	***
Time=Pretest	-0.99333	0.08851	857.00006	-11.223	< 2e-16	***
Time=Delayed.posttest	-0.76000	0.08851	857.00006	-8.587	< 2e-16	***
Group=Experimental Time=Pretest	-0.76000	0.12517	857.00006	-6.072	1.90e-09	***
Group=Experimental Time=Delayed.posttest	-0.08667	0.12517	857.00006	-0.692	0.489	

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The interaction between time and group regarding the comprehensibility of rhythm is plotted in Figure 5.20.

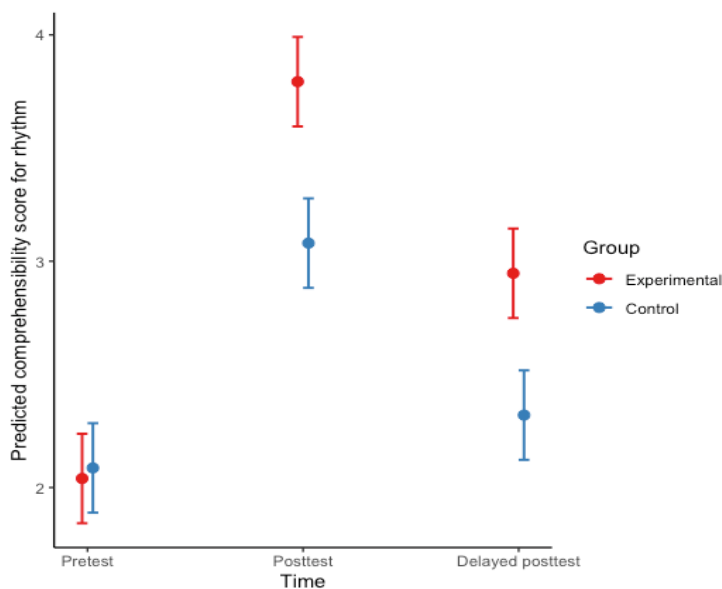


Figure 5.20. Model prediction for comprehensibility of rhythm based on time (pretest, posttests) and group (experimental, control).

### 5.5.3 Final Intonation

#### 5.5.3.1 Descriptive Results for Comprehensibility of Final Intonation

This section discusses the comprehensibility of final intonation in the experimental and control groups (see Table 5.31).

Table 5.31: Descriptive Statistical Summary for Comprehensibility of Final Intonation

Group	Test					
	Pretest		Immediate Posttest		Delayed Posttest	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Experimental	2.50	1.00	7.20	0.75	7.20	0.82
Control	2.43	0.95	4.18	1.31	3.48	1.03

The mean score of the experimental group in the pretest ( $M = 2.50$ ) was close to that of the control group ( $M = 2.43$ ), suggesting that any difference in the posttests might be a result of the intervention. In the immediate posttests, the mean score of the experimental group ( $M = 7.20$ ) was higher than the control group ( $M = 4.18$ ). This suggested the explicit intervention was effective. The experimental group attained the same score in the delayed posttest ( $M = 7.20$ ), while the control group’s mean score dropped ( $M = 3.48$ ). Figure 5.21 illustrates how these scores compare.

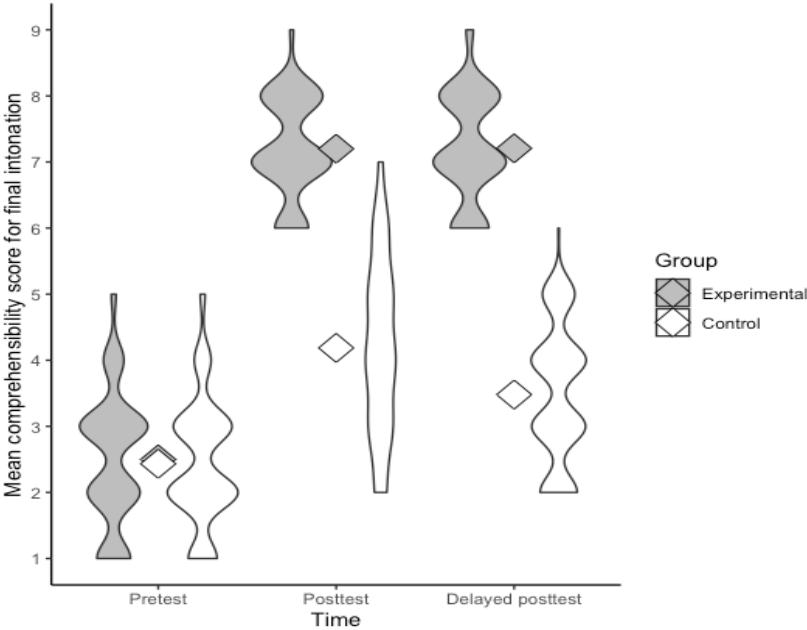


Figure 5.21. Mean comprehensibility scores of final intonation.

**5.5.3.2 Modelling Results for Comprehensibility of Final Intonation**

A linear mixed effects model was fit in R using the same formula as the full model for comprehensibility of segmentals, lexical stress, and rhythm. For the intercept (experimental

group and pretest), the predicted comprehensibility score was 2.43333. Similar to segmental and suprasegmental features for comprehensibility, time, group, and their two-way interaction were significant predictors of the comprehensibility of final intonation. Table 5.32 presents the final model.

The experimental group was not significantly different from the control group in the pretest, indicating that any difference in the posttests could be a result of the intervention. The experimental group was significantly more likely to be more comprehensible in the immediate and delayed posttests than the pretest, the control group less so.

Table 5.32: Model for Comprehensibility of Final Intonation with Pretest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	2.43333	0.15311	39.43388	15.892	<2e-16	***
Group=Experimental	0.06667	0.21398	38.16494	0.312	0.757	
Time=Posttest	1.75333	0.09940	856.99993	17.639	<2e-16	***
Time=Delayed.posttest	1.04667	0.09940	856.99993	10.530	<2e-16	***
Group=Experimental Time=Posttest	2.94667	0.14057	856.99993	20.962	<2e-16	***
Group=Experimental Time=Delayed.posttest	3.66000	0.14057	856.99993	26.036	<2e-16	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

With the immediate posttest as the reference level, the predicted comprehensibility score was 4.1867 (see Table 5.33).

Table 5.33: Model for Comprehensibility of Final Intonation with Posttest as Reference Level

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	4.1867	0.1531	39.4339	27.344	< 2e-16	***
Group=Experimental	3.0133	0.2140	38.1649	14.082	< 2e-16	***
Time=Pretest	-1.7533	0.0994	856.9999	-17.639	< 2e-16	***
Time=Delayed.posttest	-0.7067	0.0994	856.9999	-7.109	2.46e-12	***
Group=Experimental Time=Pretest	-2.9467	0.1406	856.9999	-20.962	< 2e-16	***
Group=Experimental Time=Delayed.posttest	0.7133	0.1406	856.9999	5.074	4.77e-07	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The experimental group was judged to be significantly more comprehensible than the control group. However, the difference with the delayed posttest did not reach a significant

level for the experimental group, and the control group showed an even smaller difference with the pretest. Furthermore, there was additional regression in the delayed posttest. The interaction between time and group in terms of final intonation is visualised in Figure 5.22.

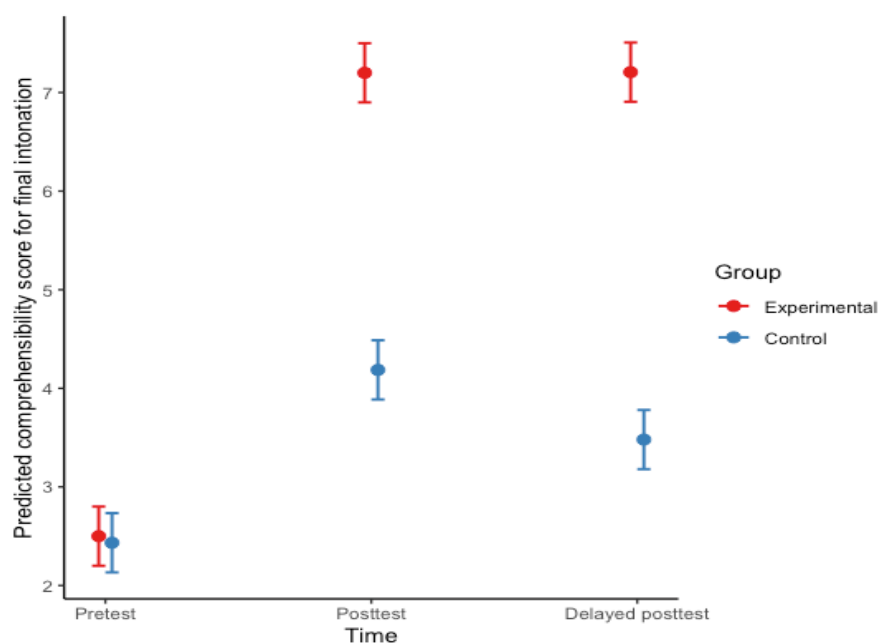


Figure 5.22. Model prediction for comprehensibility of final intonation based on time (pretest, posttests) and group (experimental, control).

## 5.6 L2 Motivation as Related to Pronunciation Improvements

This section reports the descriptive and modelling results of the four motivational dimensions: intended effort, ideal L2 self, ought-to L2 self, and attitudes toward learning English. The section explores potential relationships between motivational dimensions and group (segmental, suprasegmental, and control) to see whether the experimental groups—which were more intelligible and comprehensible than the control group across the immediate and delayed posttests—were also more motivated and what type of motivation they had.

### 5.6.1 Descriptive Results

The mean scores of all groups were compared across the four motivational dimensions, as shown in Table 5.34.

Table 5.34: Descriptive Statistical Summary for Motivational Dimensions

Group	Motivational Dimension							
	Intended Effort		Ideal L2 Self		Ought-To L2 Self		Attitudes Toward Learning English	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Segmental	3.86	1.17	3.62	1.17	3.49	1.41	3.93	0.90
Suprasegmental	4.22	1.06	4.78	0.75	3.54	1.13	4.64	0.92
Control	3.70	1.23	4.50	0.90	3.74	1.22	4.37	0.98

In terms of intended effort, the mean motivation score of the suprasegmental group ( $M = 4.22$ ) was higher than the segmental ( $M = 3.86$ ) and control ( $M = 3.70$ ) groups (see Figure 5.23).

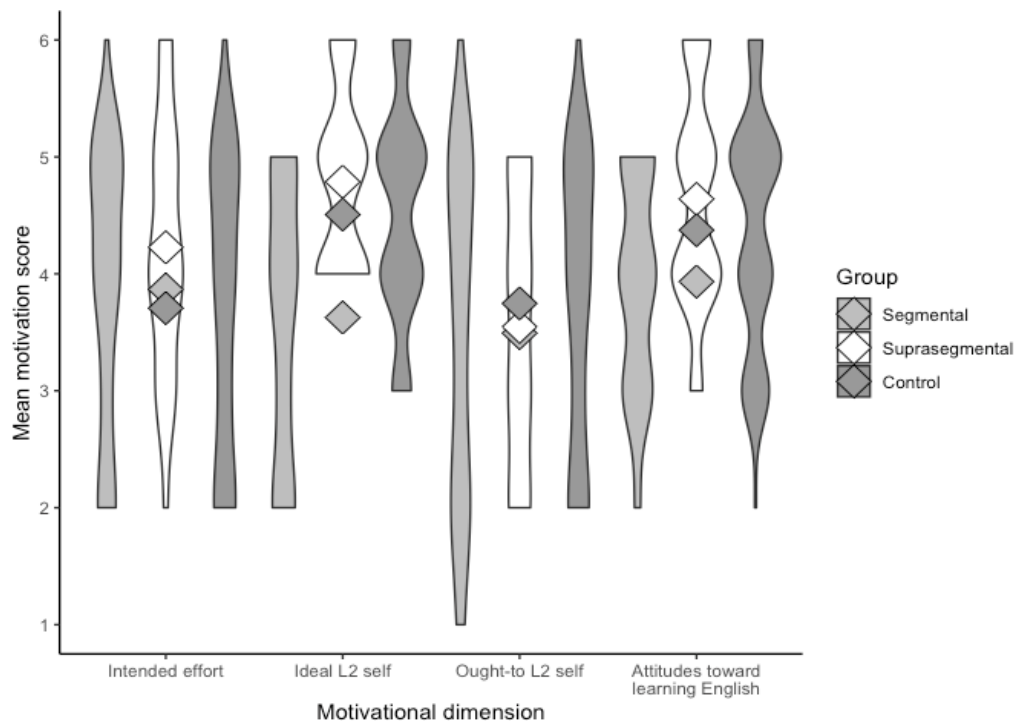


Figure 5.23. Mean scores across the four motivational dimensions.

This meant the suprasegmental group was more motivated to put effort into learning English. With ideal L2 self, the mean motivation score of the suprasegmental group ( $M = 4.78$ ) was likewise higher than the control ( $M = 4.50$ ) and segmental ( $M = 3.62$ ) groups, showing this group was more motivated to imagine themselves using English in different contexts for different purposes. Regarding ought-to L2 self, this time the mean motivation of

the control group ( $M = 3.74$ ) was higher than the segmental ( $M = 3.49$ ) and suprasegmental ( $M = 3.54$ ) groups, indicating the control group was more motivated to learn English because they viewed it as indispensable in their lives. With respect to attitudes toward learning English, the mean motivation of the suprasegmental group ( $M = 4.64$ ) was once again higher than the control ( $M = 4.37$ ) and segmental ( $M = 3.93$ ) groups. Thus, the suprasegmental group reported more positive attitudes toward learning English.

### 5.6.2 Modelling Results for Motivation

A linear mixed effects model was fit in R (see Table 5.35).

Table 5.35: Model for Motivation

Group / Dimension	Estimate	Std. Error	Df	<i>t</i>	Pr(> t )	Sig.
(Intercept)	3.7067	0.1723	40.5627	21.514	< 2e-16	***
Segmental Group	0.1600	0.1804	338.0748	0.887	0.37576	
Suprasegmental Group	0.5200	0.1804	338.0748	2.882	0.00420	**
Ideal L2 Self	0.8000	0.2376	37.6489	3.367	0.00176	**
Ought-To L2 Self	0.0400	0.2376	37.6489	0.168	0.86720	
Attitudes Toward Learning English	0.6667	0.2376	37.6489	2.806	0.00790	**
Segmental Group	-1.0400	0.2434	830.0000	-4.272	2.16e-05	***
Ideal L2 Self						
Suprasegmental Group	-0.2400	0.2434	830.0000	-0.986	0.32445	
Ideal-To L2 Self						
Segmental Group	-0.4133	0.2434	830.0000	-1.698	0.08988	.
Ought-To L2 Self						
Suprasegmental Group	-0.7200	0.2434	830.0000	-2.958	0.00319	**
Ought-To L2 Self						
Segmental Group	-0.6000	0.2434	830.0000	-2.465	0.01391	*
Attitudes Toward Learning English						
Suprasegmental Group	-0.2533	0.2434	830.0000	-1.041	0.29831	
Attitudes Toward Learning English						

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The motivation score served as a dependent variable. The fixed effects in the full model included a two-way interaction between motivational dimensions (intended effort, ideal and ought-to L2 self, and attitudes toward learning English) and group (segmental, suprasegmental, and control). Participant and question were random effects. The formula for the full model was  $\text{Score} \sim \text{Group} * \text{Motivational Dimension} + (1|\text{Participant}) + (1|\text{Question})$ . The control group and intended effort were the baseline, and the predicted motivation score was 3.7067.

Motivational dimensions, group, and the two-way interaction between them were significant predictors of motivation.

The control group had significantly lower intended effort than the segmental group, suggesting that group felt less of a need to put effort into learning English. The suprasegmental group was significantly more likely than the control group to intend to put effort into learning English. The control group was significantly more likely to imagine an ideal L2 self or express positive attitudes about learning English than to report an intended effort to learn English. However, it did not show a significance difference between ought-to L2 self and intended effort.

There were four significant interactions. The interaction between group and motivational dimension is plotted in Figure 5.24.

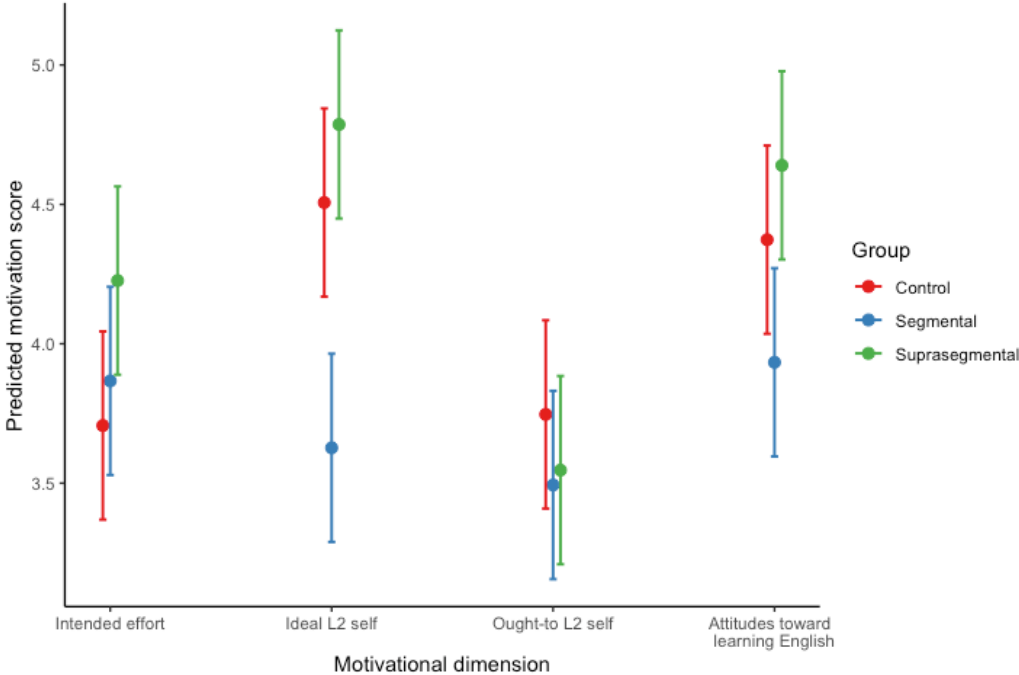


Figure 5.24. Model prediction for motivation score in the interaction between motivational dimensions and groups.

The two-way interaction between motivational dimension and group indicated that the segmental group was significantly less motivated by ideal L2 self. Unlike the suprasegmental

group, which was significantly much less motivated to learn English as an obligation (ought-to self), the segmental group was only slightly unmotivated in this regard. In addition, the segmental group had significantly more negative attitudes toward learning English than the suprasegmental group.

**5.7 How Learners Perceived the Treatment**

This section presents the results of the post-intervention questionnaire. The first part of the questionnaire was composed of two themes introduced in seven close-ended questions measured on a 4-point Likert scale ranging from 1 (“strongly disagree”) to 4 (“strongly agree”). The aim was to determine the segmental and suprasegmental groups’ overall satisfaction with the intervention and its outcomes.

The second part of the questionnaire consisted of three open-ended questions. The first and second asked what participants found were the most or least beneficial components of the pronunciation treatment. The third question explored what amendments the participants would suggest in order to make the treatment more beneficial. The following subsections present the results of the post-intervention questionnaire.

**5.7.1 Descriptive Results of the Closed-Ended Questions**

The mean scores of the segmental and suprasegmental groups were compared against the two themes, as displayed in Table 5.36.

Table 5.36: Descriptive Statistical Summary for Post-Intervention Theme Score

Group	Questionnaire Theme			
	Intervention Satisfaction		Intervention Outcomes	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Segmental	2.01	0.85	1.93	0.83
Suprasegmental	3.50	0.56	3.28	0.78

The segmental group’s mean satisfaction score ( $M = 2.01$ ) was much lower than the suprasegmental group ( $M = 3.50$ ). This revealed that the suprasegmental group was more

satisfied with the intervention. With respect to intervention outcomes, the mean score of the suprasegmental group ( $M = 3.28$ ) was much higher than the segmental group ( $M = 1.93$ ). In other words, the suprasegmental group was more likely to consider the intervention a success.

Figure 5.25 compares these scores.

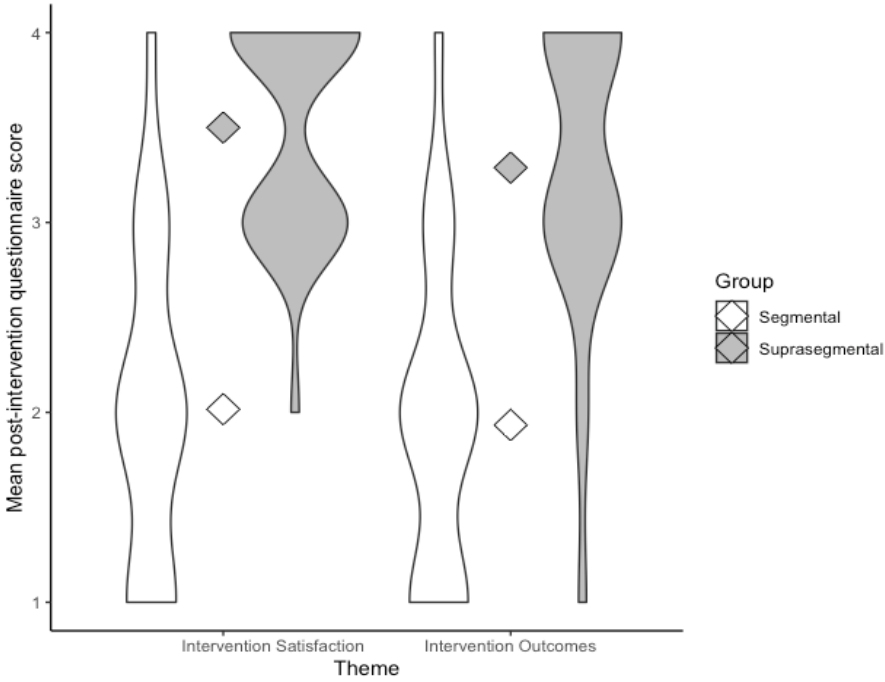


Figure 5.25. Post-intervention mean scores for satisfaction and outcomes.

**5.7.2 Modelling Results of the Closed-Ended Questions**

A linear mixed effects model was fit in R. The score served as a dependent variable. The fixed effects in the full model included a two-way interaction between theme (intervention satisfaction, intervention outcomes) and group (segmental, suprasegmental). Participant and question were random effects. The formula for the full model of the post-intervention questionnaire was  $Score \sim Group * Theme + (1|Participant) + (1|Question)$ . Table 5.37 presents the final model.

Table 5.37: Model for the Theme of the Post-Intervention Questionnaire

	Estimate	Std. Error	Df	<i>t</i>	Pr(>  <i>t</i>  )	Sig.
(Intercept)	1.98095	0.07454	208.00000	26.57	<2e-16	***
Suprasegmental Group	1.42857	0.10542	208.00000	13.55	<2e-16	***

Note.  $p < 0.1^*$ ,  $p < 0.05^{**}$ ,  $p < 0.01^{***}$ ,  $p < 0.001$ .

The segmental group and intervention satisfaction were the reference levels. The predicted post-intervention questionnaire score was 1.98095. Group was found to be a significant predictor. The suprasegmental group was significantly more satisfied with the intervention than the segmental group. No significant two-way interactions were detected between group and theme, as illustrated in Figure 5.26.

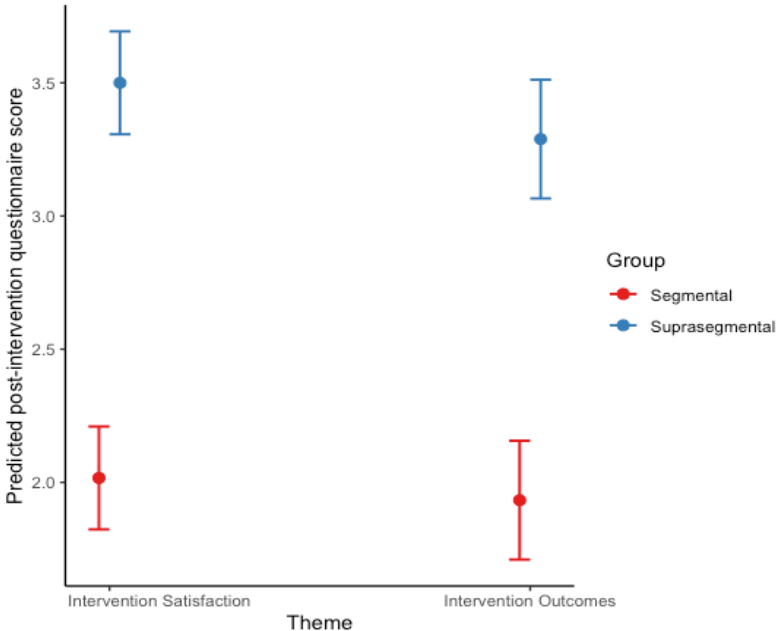


Figure 5.26. Model prediction for the interaction between theme (intervention satisfaction, outcomes) and group (segmental, suprasegmental).

**5.7.3 Narrative Responses for the Open-Ended Questions**

This section gives an account of the segmental and suprasegmental groups’ narrative responses qualitatively collected from three open-ended questions. The main themes of each question were extracted and then categorised.

Regarding the first question, the most beneficial components in the explicit intervention, participants' narrative responses were categorised into four themes. Fourteen segmental and suprasegmental participants expressed that the practices during the intervention were more beneficial, and 22 indicated the activities helped make their speech more intelligible and comprehensible. Twelve found the experienced native English-speaking teacher more helpful for introducing the sessions, while 22 considered the materials more effective as they played a crucial role in explaining the segmental and suprasegmental features that were taught.

With respect to the second question, regarding the least beneficial components of the explicit intervention, two themes emerged from participants' narrative responses. In the first, a large majority of segmental and suprasegmental participants stated that the duration of the sessions was insufficient. In the second, 10 participants claimed that the opportunities for individual practice were less effective.

The third question, regarding any changes participants would suggest making to the treatment, produced three themes. First, 17 participants proposed assigning more time for instruction sessions. Second, 13 participants suggested technology could be used to make delivery more effective and help students learn at a time and place convenient for them. Third, 11 participants suggested increasing the number of opportunities for individual practice.

## **5.8 Raters' Perceptions of the Pronunciation of Participants**

This section presents the results of the post-assessment questionnaire. It includes two parts: the raters' perceptions of the intelligibility of the audio-recordings and their perceptions of the recordings' comprehensibility.

### **5.8.1 Perceived Intelligibility of Segmental and Suprasegmental Features**

Three themes emerged from the raters' narrative responses about intelligibility. The first focused on the difficulty of differentiating between words that only differ by one sound

(minimal pairs), particularly at the beginning and end of a word. For instance, one rater pointed out that “it would be more difficult for me to determine the difference between ‘sung’ and ‘sunk’ because to me, they sounded very similar.” Another rater mentioned that “the ‘sh’ and ‘ch’ were difficult to determine because their pronunciation was very similar. I think out of all the ones I did, this one was the most difficult.” Along the same lines, another rater stated that “Some words [were] said incorrectly due to the wrong pronunciation of specific phonemes.” These responses revealed the importance of differentiating between each segment in individual words to enhance intelligibility, especially with minimal pairs. Hence, segments emerged as a key factor in distinguishing words.

The second theme highlighted the issue of stress misplacement as some participants, according to the raters, were not able to correctly place stress. For instance, one rater stated that “Some words [were] pronounced incorrectly due to the incorrect stress placement.” Another made this comment: “Mumbled. Incorrect pronunciation of syllables for individual words.” Raters indicated that stress placement plays a key role in determining how intelligible the participants were, particularly in individual words.

The third theme shed light on rhythm, as some sentences were less intelligible for the raters because of the lack of correct stress placement within a sentence. For example, one rater stated that “the rhythm of phrases sometimes made it hard to understand individual words.” To emphasise the significance of rhythmic patterns across the pretest and posttests, another rater drew attention to the “Lack of awareness of difference between sentences based on rhythm.” These comments illustrated that rhythm was a major factor impeding students from being more intelligible.

### 5.8.2 Perceived Comprehensibility of Segmental and Suprasegmental Features

Three themes emerged from the raters' narrative responses about comprehensibility. The first was that individual segments were not a substantial factor making students more comprehensible compared to intelligibility. This might be a result of the raters relying on context to understand a sentence as a whole. For instance, one rater stated, "Also, some words were just mispronounced incorrectly like 'put' sometimes sounded like 'but' and 'leave' sounded like something else. However, I ultimately comprehended the whole sentence." Another rater stated that "although individual segments sometimes were produced wrongly, the speaker was still comprehensible." This finding revealed that raters assumed that individual segments did not affect comprehensibility as other features—such as stress, rhythm, and intonation—could be more crucial to understanding L2 speakers' English.

The second theme was the importance of stress and rhythm patterns in producing comprehensible speech. As one rater noted, "Incorrect stress and wrongly pronounced words also made it difficult to comprehend", while another stated that "I found the 'volunteer' sentence difficult to hear a lot of the time, and also the sentence to do with a 'permit'. Also the 'university record' was hard to hear a lot of the time as well." In the same vein, another rater stated, "Those that were more comprehensible were closer to native pronunciation, stress, and rhythm." Another rater stated, "I also think the stress of certain parts of the sentence were different to a native speaker, and that also made it more difficult to follow." As shown in these comments, the raters emphasised the importance of stress and rhythm to increase comprehensibility in L2 English speech.

The third theme was intonation as a contributing factor to comprehensible speech. One rater described using intonation as a key factor to understand student recordings: "Where there was native-speaker-like intonation, this helped understanding." This issue was apparent when

trying to determine whether a sentence was a question or a statement, as one rater found that “Some sentences could have been questions; i.e., there was a vague rising intonation; also, no context.” Another rater mentioned that “intonation was not at all used as a key to differentiate between given sentences.” These comments demonstrated the importance of intonation patterns for attaining comprehensible speech in English.

## **5.9 Chapter Summary**

The segmental group’s results from the immediate posttest revealed that intelligibility and comprehensibility improved significantly over the pretest and control group. However, the group failed to maintain this improvement over time, except for the segment /v/. Regarding the suprasegmental group, the immediate posttest showed significant improvement in intelligibility and comprehensibility over the pretest and control group. In addition, the suprasegmental group showed intelligibility and comprehensibility improvement over time, with the exception of rhythmic patterns for intelligibility and lexical stress for comprehensibility. Motivation was revealed to be related to L2 pronunciation development in the suprasegmental group. The suprasegmental group’s responses indicated the explicit pronunciation instruction was effective. Native speakers’ comments showed that intelligibility and comprehensibility could be impeded by a wide range of segmental and suprasegmental challenges. The next chapter discusses the results according to the research questions and the literature.

## **Chapter 6: Discussion**

### **6.1 Introduction**

This chapter discusses the results of the study in light of the research questions and relevant literature. The study examined the effectiveness of explicit pronunciation instruction at helping Saudi beginning-level EFL learners attain and maintain more intelligible and comprehensible speech. The chapter is divided into nine sections, six of which are based on the six research questions.

The next section summarises the major results. The third section discusses the results of the first research question in relation to the effectiveness of explicit instruction on improving the intelligibility and comprehensibility of segmentals. The fourth section is devoted to the results of the second research question, focusing on the effect of explicit instruction on suprasegmentals. Addressing the third the research question, the fifth section discusses the ability of the Saudi beginning-level EFL learners to retain pronunciation gains over time. The sixth section focuses on the results of the fourth research question, regarding L2 motivation and attitudes in English learning. The seventh section discusses the results of the fifth research question, regarding participants' perceptions about the explicit intervention. The eighth section explains the results of the sixth research question, exploring the raters' assessment of the intelligibility and comprehensibility of segmental and suprasegmental features. The ninth section summarises the main components of the chapter.

### **6.2 Summary of Major Findings**

Table 6.1 gives an overview of the intelligibility and comprehensibility results for segmental and suprasegmental features. The control group was significantly more intelligible than the experimental group in all segmental features in the pretest. Despite that, the experimental groups showed significantly more improvement than the control group in the

immediate posttest, except for the segment /f/. Additionally, only the experimental group results for /v/ showed significant improvement over the control group in the delayed posttest. With respect to segmental comprehensibility, the groups were not significantly different in the pretest. However, the experimental groups in the immediate posttest were significantly more likely to be intelligible than the control group. Neither group had delayed posttest results significantly different from the immediate posttest.

Table 6.1: A Summary of Segmental and Suprasegmental Results across a Battery of Tests

Group	Construct	Feature	Significant improvement over control group		
			Pretest	Immediate Posttest	Delayed Posttest
Segmental	Intelligibility	/p/	✗	✓	✗
		/v/	✗	✓	✓
		/f/	✗	✗	✗
		/ŋ/	✗	✓	✗
	Comprehensibility	All four segmentals	✗	✓	✗
Suprasegmental	Intelligibility	Lexical stress	✗	✓	✓
	Comprehensibility	Lexical stress	✗	✓	✗
	Intelligibility	Rhythm	✗	✓	✗
	Comprehensibility	Rhythm	✗	✓	✓
	Intelligibility	Final intonation	✗	✓	✓
	Comprehensibility	Final intonation	✗	✓	✓

In the pretest suprasegmental results for lexical stress, the experimental group scored significantly lower than the control group. However, the experimental group scored significantly higher than the control group in the immediate posttest. In the delayed posttest, the experimental group scored significantly higher than the control group in terms of lexical intelligibility. With respect to rhythm, the experimental groups were not significantly different from the control group in the pretest but were significantly more intelligible and comprehensible than the control group in the immediate posttest. The experimental groups scored significantly higher than the control group in the delayed posttest only in terms of comprehensibility. In the suprasegmental pretest results for final intonation, the experimental

groups were not significantly different from the control group in terms of intelligibility or comprehensibility. However, they were significantly more likely to be intelligible and comprehensible than the control group in the immediate and delayed posttests.

The suprasegmental group had significantly higher intended effort and ideal L2 self than the segmental and control groups. The segmental group, in contrast, was not significantly more motivated than the control group in terms of intended effort and was significantly less motivated than the control group in terms of ideal L2 self. Regarding ought-to L2 self, the control group was significantly more motivated to learn English as a sign of necessity than the experimental groups. Regarding attitudes toward learning English, the suprasegmental group was significantly more motivated than the other two groups, while the control group was more motivated to have a positive attitude toward English than the segmental group.

Based on the quantitative data, the suprasegmental group was significantly more satisfied with the intervention and its outcomes than the segmental group. According to the qualitative data, group practice, activities, materials, and having an experienced native English-speaking teacher were the most beneficial factors in the explicit pronunciation intervention. The least beneficial factors were the duration of the intervention and not having more individual practice. Participants suggested allocating more time, using digital technology more extensively, and having more individual practice.

The raters' intelligibility assessments revealed that differentiating between minimal pairs was problematic. Misplaced stress and incorrect rhythmic patterns had a significant negative effect on intelligible speech. With respect to comprehensibility, individual sounds did not have a major negative influence. Sentence stress and intonation patterns helped make learners more comprehensible.

### **6.3 Research Question 1: Immediate Effect on Segmentals**

This section discusses the results of the first research question, which was concerned with the immediate effects of the explicit pronunciation instruction on segmental intelligibility and comprehensibility. The pretest showed a high frequency of mispronunciation of segmentals, with the segmental and control groups showing a similar number of errors. This result made it easier to determine the effectiveness of the intervention and was anticipated for several reasons.

One potential reason was the absence of the target English sounds in the segmental stimuli from the Arabic phonemic inventory, which could make it difficult for Saudi beginning-level EFL learners to produce these sounds. Learners would likely attempt to find Arabic equivalents. Thus, the Arabic sound system can play an influential role in determining the accuracy of target sounds, particularly in the initial stages of L2 learning. The issue of L1 interference in the early stages of learning has been intensively discussed in the arena of L2 speech perception and production (e.g., Best, 1995; Best & Tyler, 2007; Flege, 1995; Flege & Bohn, 2021). The SLM (Flege, 1995) and PAM (Best, 1995) have asserted that L2 learners copy and transfer the characteristics of their L1 sound system to perceive and produce L2 sounds. This occurs as a potential consequence of categorising and assimilating L1 sounds into L2 sounds, because learners assume the L1 and L2 sounds have the same integrated phonemic system.

Another possible explanation for the pretest results might be that the learners had insufficient knowledge of English phonetics and phonology. Thus, L2 input could play a key role in determining learner awareness of L2 linguistic characteristics. A great deal of previous empirical research has placed substantial emphasis on the amount of L2 input in perceiving and producing L2 sounds (e.g., Flege, 2009, 2018).

After explicit instruction, the segmental group showed significant improvement in segmental intelligibility and comprehensibility in the immediate posttest. This result agrees with previous research on the positive immediate effects of pronunciation instruction on the intelligibility and comprehensibility of L2 speech (see Derwing, 2018).

The positive effect of explicit instruction on segmentals could be attributed to the focus and type of instruction. Such findings would lend support to Zielinski and Yates' (2014) proposed systematic approach to implement pronunciation instruction with beginning-level L2 learners. They claimed this approach would make learners more intelligible and comprehensible when pronunciation teaching went through developmental stages, such as from controlled to uncontrolled activities.

However, any interpretation of the posttest findings for segmentals must be made with caution. The segment /tʃ/ was not significantly different from the pretest like /p/, /v/, and /ŋ/ compared to the control group. Nevertheless, the segmental group's posttest results for /tʃ/ were significantly different from its pretest results. There are a number of potential explanations for these results. The first is cognitive load, as participants were taught /tʃ/ in the last week of the treatment, meaning they might have been more tired and less focused. Second, the targeted segmental stimuli were presented in isolated words in a carrier sentence. Third, the nature of the learning context must not be ignored, as the participants were only required to be exposed to and interact in English in the classroom. Consequently, lack of exposure to English could have prevented learners from increasing their awareness of the phonetic information of English sounds. Production could also have been influenced by phonological environment, i.e., whether the target segment appeared in the onset or coda of words. Another possible explanation was the nature of the forced identification task used to assess the intelligibility of participants'

segmental production. Since the raters only had two options, this might have caused greater cognitive demand and presented them with a lack of better choices.

#### **6.4 Research Question 2: Immediate Effect on Suprasegmentals**

This section discusses the results of the second research question, which assessed the intelligibility and comprehensibility of suprasegmental features, including lexical stress and intonation. The section offers various potential reasons for why the participants failed to improve the intelligibility and comprehensibility of their English rhythm in the immediate posttest.

As expected, learners' pretest suprasegmental results showed relatively low comprehensibility and intelligibility, likely due to a lack of knowledge about English prosody. To make up for this gap, learners tended to transfer their Arabic prosodic rules to their English production in the pretest. As a practical example, English is a stress-timed language, while Arabic is syllable-timed (Celce-Murcia et al., 2010). This difference could have made it harder for the learners to place stress on the correct syllable in English. This, in turn, can lead to a loss of communication, as Levis (2018) stated that misplaced word stress in English might interfere with spoken communication.

In addition, English rhythm is used to divide up continuous speech by native listeners in order to differentiate between actual words and unactual words that should be excluded (Levis, 2018, p. 136). Such segmentation strategies are not present in Arabic; as a result, it could be difficult for Saudi beginning-level EFL learners to correctly segment English discourse (cf. Levis, 2018). Intonational structure in L2 speech perception and production has received considerable attention as it plays a crucial role in establishing informational and social convergence (e.g., Pickering, 2009, 2012, 2018; Pickering & Litzenberg, 2011). Overall,

intelligibility and comprehensibility of L2 English speakers are more likely to be negatively affected when prosodic features are used incorrectly.

The pretest findings for suprasegmentals support previous L2 pronunciation research about the intelligibility and comprehensibility of L2 beginning-level speakers prior to an explicit intervention. As a result, such difficulties might lead to loss of confidence and willingness to communicate in English (Zielinski & Yates, 2014).

The suprasegmental immediate posttest results showed significant improvements in lexical stress and intonation. This result is consistent with a great deal of L2 pronunciation research that has found explicit pronunciation instruction to be more effective with suprasegmental features. For instance, in Gordon et al.'s (2013) study, the suprasegmental group showed significant pronunciation improvements in intelligibility and comprehensibility compared to the control group.

The suprasegmental results of the present study for comprehensibility are also in line with Gordon and Darcy (2016). They conducted a short-term experimental study to investigate the development of comprehensible speech among ESL learners. The suprasegmental group received explicit instruction, whereas the control group had non-explicit instruction. The suprasegmental group ultimately showed greater improvement than the control group.

To understand why learners improved their lexical stress and intonation patterns after the intervention but failed to do so with rhythm, a number of possible explanations can be considered. First, much prior research has emphasised that rhythmic patterns are far more difficult to teach L2 learners, especially beginners (Levis, 2018). Second, this finding might be attributed to the method used to measure rhythm intelligibility. A transcription task was used, but comprehensibility did not improve even when measured in the same way as other suprasegmental features. The task required placing stress on content words, and the correct

stressed words were then counted, while lexical stress and intonation were assessed by applying a two-way forced choice identification task as the assessors were asked to choose between two options.

There were cases where the experimental and control groups showed a significant within-group improvement in the immediate posttest. For instance, the control group significantly improved in both posttests in terms of lexical stress intelligibility and comprehensibility. However, the experimental group showed a greater significant improvement than the control group in lexical stress intelligibility and comprehensibility in the immediate posttest and in intelligibility in the delayed posttest. This could indicate that non-explicit instruction could improve beginners' pronunciation but to a lesser extent than explicit instruction. This result is in line with Derwing and Munro (2015):

In fact, learners' phonological development seems to be most active during the first six months to a year of intensive language training and/or exposure to the L2.... there appear to be a window of maximal opportunity to develop L2 phonological skills that will result in clear. (p. 99)

### **6.5 Research Question 3: Long-Term Effectiveness of Treatment**

One of the primary objectives of this study was to identify long-term positive effects from explicit pronunciation instruction. For this purpose, the delayed posttest results were statistically compared to the pretest within each experimental group (segmental and suprasegmental). Another comparison was performed between each experimental group and the control group in the delayed posttest to look for significant differences due to the explicit instruction.

In the delayed posttest, the suprasegmental group was more successful in retaining its suprasegmental gains in terms of intelligibility (except rhythm) and comprehensibility (except

lexical stress) compared to the control group. Although the segmental group did not show significant improvement over the control group in the delayed posttest (except for the segment /v/), a within-group comparison for the segmental group did show significant improvement over time. These findings offer evidence of the positive effect that explicit instruction can have on pronunciation over time regardless of learner proficiency level. Furthermore, the findings of the delayed posttest respond to previous suggestions to implement explicit pronunciation instruction on a long-term basis (e.g., Derwing & Munro, 2015).

The results of the delayed posttest for /v/ intelligibility and suprasegmental features (except lexical stress for comprehensibility and rhythm for intelligibility) align with what has been demonstrated in recent empirical research that implemented a delayed posttest and revealed positive effects of pronunciation instruction. For instance, Lee et al. (2020) conducted a pretest, posttest, and delayed posttest using a perception- or production-based approach in a two-week treatment on English segmentals or suprasegmentals. In all treatment groups, the participants' pronunciation became substantially more accurate in both posttests. Although the duration of instruction was short, the delayed posttest showed that pronunciation instruction with a perception-based focus tended to be more effective with segmental and suprasegmental features.

Zhang and Yuan's (2020) study is another example of the long-term benefits of pronunciation instruction in line with the delayed posttest results of this study. They investigated how the English pronunciation of second-year Chinese undergraduate students could be developed in an 18-week period of instruction by comparing the effects of segmental and suprasegmental instruction. The results of both groups showed significant improvement in the immediate posttest, but only the suprasegmental group retained those gains by the time of

the delayed posttest. This demonstrates how the positive effects of pronunciation instruction can last long after formal instruction is over.

Couper's (2006) classroom-based study explored whether the phonological competence of high-intermediate English learners in New Zealand could be retained over time after a two-week explicit pronunciation intervention. Although the study aimed to measure accentedness rather than intelligibility or comprehensibility, the results of the delayed posttest showed that learners were still significantly better than before the intervention in terms of avoiding vowel epenthesis. This is once again in alignment with the empirical evidence of the current study about the positive effects of pronunciation instruction over time.

Before explaining why some segmental and suprasegmental features failed to show significant improvement in the delayed posttest, it is first important to acknowledge that much L2 pronunciation research with a delayed posttest phase has been conducted in English-speaking contexts (e.g., Alameen, 2014), in different EFL contexts than the present study (e.g., Zhang & Yuan, 2020), on segmental and suprasegmental features for Arabic-speaking EFL learners that were different from those in this study (e.g., Evans & Alshangiti, 2018), and in a laboratory setting (e.g., Baese-Berk, 2019). These differences make it difficult to compare the delayed posttest results of this study with other empirical L2 pronunciation research.

Despite these differences, there are a number of possible explanations for the lack of positive long-term effects on some segmental and suprasegmental features. Major factors include the short duration of the explicit pronunciation instruction, participants' low level of English proficiency (as beginning-level learners), and the relatively small sample size. In addition, individual differences (e.g., motivation and willingness to communicate) and a lack of compulsory exposure to English outside the classroom represent other factors that might have inhibited long-term retention.

Regarding duration, previous research has strongly suggested extending the duration of explicit pronunciation instruction, because a shorter treatment might have less of a positive influence on long-term memory. For instance, Thomson and Derwing (2014) pointed out that intelligibility and comprehensibility improvements require a longer duration of pronunciation instruction. Thus, it would be optimal to offer several weeks or months of instruction, not just a few hours or days (p. 336). Similarly, Ruellot (2011) implemented a two-hour training session for intermediate learners of French in order to improve their vowel production. The delayed posttest revealed learners just one week later failed to retain their improvements and returned to their pre-training status. This finding illustrates the importance of providing longer instruction to achieve longer-lasting gains.

Since participants were beginning-level learners, some segmental and suprasegmental features might have been more difficult for them to retain. A large body of L2 pronunciation literature has discussed learner proficiency level as a factor in segmental and suprasegmental acquisition (e.g., Saito, 2018). Nevertheless, the delayed posttest findings for segment /v/ are in line with research that has found pronunciation instruction can be tremendously important even if improvements are limited at the initial stages of L2 phonology acquisition. For example, Zielinski and Yates (2014) emphasised that “pronunciation can and should be taught in beginning-level classes and can be usefully integrated into the curriculum at that level” (p. 65).

Individual differences between L2 learners are a third possible explanation. Such an essential factor, however, has received little attention in the area of L2 speech production and perception acquisition. Edwards (2018) conceptualised a number of variables that might link L2 pronunciation and individual differences, such as aptitude, motivation, willingness to communicate, and anxiety. With respect to the role of individual differences in the current study, the suprasegmental group was more likely to be motivated across different motivational profiles

(the four motivational profiles are discussed in the next section). This might indicate that individual differences could be a significant predictor of the suprasegmental group's better performance in the delayed posttest compared to the segmental group.

The delayed posttest results show a close alignment with other research about individual differences and L2 pronunciation. In an empirical study on the relationship between cognitive ability and individual differences in L2 phonological acquisition, Darcy et al. (2015) found that Korean learners of English who showed a higher working memory developed more native-like pronunciation. In addition, Saito (2019) investigated the role of awareness as a predictor of individual differences in the development of English /ɪ/ among Japanese EFL learners. The findings indicated that learner awareness played a significant role in pronunciation accuracy.

Limited exposure to English outside the classroom is another factor that might explain why the segmental group failed to attain more intelligible and comprehensible speech than the control group in the delayed posttest. The literature postulates that pronunciation achievement can be positively influenced by the amount of exposure to the target language (e.g., Grant, 2014b). Supporting this claim, Flege (2009) stated that exposure to English should help L2 learners show pronunciation improvement regardless of whether they interact with native or non-native speakers.

#### **6.6 Research Question 4: Motivations and Attitudes as Predictors**

This section discusses the results of the fourth research question, relating to the effects of motivation and attitudes on pronunciation. With respect to the first motivational profile, intended effort, the suprasegmental group was more motivated to put effort into learning English than the other two groups. This result is evidence of the relationship between L2 motivation and pronunciation development, as the suprasegmental group also showed better immediate and long-term gains in intelligibility and comprehensibility than the other groups.

In line with the recent literature on L2 motivation and its relationship to L2 pronunciation development, this study's results for the "intended learning effort" motivational profile confirmed the importance of putting effort into L2 learning to improve oral proficiency. In an empirical longitudinal study of the relationship between motivation and L2 speech development, Saito et al. (2017) found that comprehensibility significantly improved among 40 first-year university Japanese EFL learners during one semester. Furthermore, these learners were more likely to show a high level of motivation to learn English and put in more effort to reach their goals. Thus, they were more motivated in terms of intended learning effort and ideal L2 self. Similarly, the suprasegmental group in the present study significantly improved when they put more effort into learning English. Such effort might be effective even without a direct intensive focus on teaching pronunciation (Edwards, 2018).

In terms of the second motivational profile, ideal L2 self, the suprasegmental group was significantly more motivated than the other two groups. As that group also had more intelligible and comprehensible speech in the posttests, this result clearly shows the importance of how L2 learners imagine themselves in the future, which can foster their ability to produce clear, efficient, and effective communication. These findings support previous research. For example, Huensch and Thompson (2017) found a positive correlation between the importance of improving L2 pronunciation and how L2 learners imagined themselves (ideal L2 self). As another example, Alshehri (2009) found a relationship between how L2 learners visualised themselves and their interaction (including spoken interaction) in the target language. Learners with a high level of motivation in terms of ideal L2 self tended to naturally exhibit more positive behaviour, envisioning themselves interacting in settings where the target language was mainly used.

The third motivational dimension was ought-to L2 self. Although the control group was slightly more motivated to learn and use English as a necessity than the experimental groups, the results showed no significant improvements in intelligible and comprehensible speech for the control group from the pretest to the posttests or in comparison with the experimental groups. This lack of an effect from ought to L2 self is in line with Nagle (2018). Nagle's empirical study explored to what extent motivation in L2 learning played a role in developing L2 pronunciation across three academic semesters among English-speaking learners of Spanish. Ought-to L2 self was found to be less effective than other motivational profiles.

A large body of literature has called for researchers to re-examine some theoretical aspects of ought-to L2 self, as it does not provide a detailed description of L2 learners' perceptions about their learning responsibilities (e.g., Teimouri, 2016). As a consequence, the results of the present study for this motivational profile must be interpreted with caution as the control group did not become more intelligible and comprehensible than the experimental groups despite being more likely to view learning English as necessary due to personal, academic, familial, and social reasons.

In the results for the fourth motivational profile, the suprasegmental group had a more positive attitude toward learning English than the other groups, while the control group was more motivated in this regard than the segmental group. As the suprasegmental group had better pronunciation scores in the posttests than in the pretest, this could offer further evidence that positive attitudes toward learning an L2 can improve L2 pronunciation. This finding is consistent with previous empirical research on the relationship between learner attitudes and L2 pronunciation improvement. For instance, Huensch and Thompson (2017) reported that positive attitudes were closely related to a desire to improve L2 pronunciation. Moyer (1999)

posited that late L2 learners might be able to attain L2 phonological patterns once they show positive attitudes toward an L2, including the place, type, and amount of L2 instruction.

As another possible explanation for the importance of positive attitudes toward learning the target language, L2 learners might be more intelligible and comprehensible once they become more confident in their pronunciation. A recent empirical study by Tsunemoto and McDonough (2020) offers support for this explanation. Regarding one of the attitudinal factors examined in their study, the findings revealed that pronunciation confidence led learners to become more competent in L2 pronunciation. Moyer (2007) reported similar findings on the importance of attitudes toward the target language, including the ability of L2 learners to improve their English and self-confidence in English.

Overall, the results for motivation and attitudes in L2 learning provided further evidence of the relationship between motivation and the development of L2 pronunciation. The suprasegmental group scored higher in different motivational profiles and outperformed the control group in intelligibility in the immediate posttest and comprehensibility in the delayed posttest.

### **6.7 Research Question 5: Learners' Perceptions of the Treatment**

The fifth research question explored how Saudi beginning-level EFL learners perceived the explicit pronunciation treatment. The first part of the post-intervention questionnaire collected quantitative data from the segmental and suprasegmental groups. In both themes of the questionnaire (intervention satisfaction and intervention outcomes), the suprasegmental group scored significantly higher than the segmental group. These findings provide additional support for the experimental results, as the suprasegmental group performed better in both of the posttests than the segmental group. This result might help explain the suprasegmental group's higher satisfaction with the intervention and its outcomes.

However, the quantitative findings of the post-intervention questionnaire should be interpreted with caution as the suprasegmental group failed to show improvement in rhythmic patterns in the posttests. Although the suprasegmental group's responses on the questionnaire revealed significant positive results about the intervention's effectiveness, the number of questionnaire themes were limited. In addition, the number of participants in the two experimental groups was relatively small. This line of reasoning is in line with a recent call from Saito et al. (2020):

To tackle the topic of individual differences in any aspect of L2 learning, much caution needs to be exercised in data collection and analysis. It is important to recruit a large number of participants to maintain a strong statistical power of dependent variables. (p. 22)

The qualitative part of the post-intervention questionnaire provided two additional explanations for the results of the first part. First, the pronunciation activities, having a native English-speaking teacher, and the materials used might have helped the suprasegmental group become more intelligible and comprehensible. Second, insufficient time for pronunciation instruction and a lack of individual practice might have had a negative effect on the segmental group. Although the qualitative results demonstrated explicit evidence of how effective the pronunciation instruction was, the qualitative findings should be considered more carefully, as improvements in intelligibility and comprehensibility are related to a complex variety of factors, some of which go beyond the type, duration, and content of pronunciation instruction (Lee et al., 2015; Thomson & Derwing, 2014).

### **6.8 Research Question 6: Raters' Perceptions of Learners**

This section discusses the results of the post-assessment questionnaire. The first part was about the raters' perceptions of the intelligibility of learners' segmentals and

suprasegmentals, while the second part was concerned with learners' comprehensibility. The intelligibility assessments indicated that differentiation between individual segments, misplacement of lexical stress, and incorrect use of rhythmic patterns could make L2 speech less intelligible. However, the raters did not consider individual segments to be a major obstacle to comprehensibility assessment. Stress, rhythm, and intonation patterns were substantial factors determining whether L2 learners' speech would be comprehensible to the raters.

According to the raters' comments, the experimental results of the segmental group were strongly affected by the accuracy in producing individual segments. This represents evidence of the importance of individual segments, in particular if they appear in a minimal pair and a controlled context. This finding is in keeping with a large body of L2 pronunciation literature claiming that segmentals could have a greater impact on mutual intelligibility (e.g., Bent et al., 2007; Levis, 2018; Zielinski, 2008, 2015).

Nevertheless, the raters pointed out that suprasegmental features must not be neglected when working toward greater spoken intelligibility. This recommendation provides another explanation for the suprasegmental group's failure to increase their intelligibility after receiving explicit instruction on English rhythmic patterns. Prior research has similarly revealed suprasegmentals to be crucial to intelligibility. For instance, Levis (2018) postulated that rhythmic differences appeared to indirectly contribute to a loss of intelligibility.

The raters stated that individual segments were not problematic for comprehensibility. This claim supported the segmental results for comprehensibility, as the segmental group did better than the control group in the posttests. In addition, the segmental group's delayed posttest results were better than its pretest results. These findings conflict with some research suggesting individual segments are important to comprehensibility (e.g., Saito, 2011; Suzukida & Saito, 2019; Zielinski, 2015). A possible explanation for this outcome is that the target segments that

the present study used to measure comprehensibility were presented in high-frequency words within shorter, more structured sentences. This was done because the participants were beginners. For this reason, the way segmental stimuli were presented in this study could have differed from other research. Another explanation is that raters might have relied on prosodic information rather than individual segments to rate learners' comprehensibility at the segmental level.

Regarding the importance of suprasegmental features for comprehensibility, raters assumed that stress, rhythm, and intonation patterns were crucial. This consensus highlighted how effective the instruction was, as the suprasegmental group greatly improved in terms of prosody in the posttests. Prior research has similarly found suprasegmental features to be indispensable in helping L2 learners attain more comprehensible speech (e.g., Gordon & Darcy, 2016; Kang, 2010).

## **6.9 Chapter Summary**

This chapter discussed the results in relation to the research questions and literature. The study showed generally positive effects of pronunciation instruction on the pronunciation of beginning-level EFL learners. Many factors could have played a role in this improvement, such as the effectiveness of the instruction or the type and degree of participant motivation. However, the segmental and suprasegmental groups failed to improve their production of some segmental and suprasegmental features in the posttests. A number of factors could explain this, such as the short length of instruction. The next chapter suggests pedagogical and research implications for incorporating pronunciation instruction into the beginning-level Saudi EFL classroom.

## **Chapter 7: Implications for L2 Pronunciation Pedagogy and Research**

### **7.1 Introduction**

This study offers significant pedagogical implications to help Saudi beginning-level EFL learners attain more intelligible and comprehensible pronunciation. The results of this study may also show paths for academics, instructors, policymakers, and curriculum designers. In addition, the results may help advance L2 pronunciation learning and teaching research about issues in these classrooms.

As previously noted, current research supports implementing L2 pronunciation instruction, including more detail action to make pronunciation teaching and learning more beneficial in the classroom (e.g., Levis, 2005; Derwing & Munro, 2015; Brinton, 2018b). This chapter engages with existing literature to devise avenues for improving instruction in the Saudi context.

This chapter outlines two types of implications to integrate pronunciation teaching into the beginning-level classroom. First, it focuses on how segmental and suprasegmental pronunciation can be taught, how improvements can be attained over time, how motivation can be addressed, and what learners' perceptions suggest about the effectiveness of pronunciation instruction in the Saudi beginning-level EFL classroom. Second, it discusses further research avenues and questions about pronunciation instruction in the ESL/EFL classroom in general.

### **7.2 Implications for L2 Pronunciation Pedagogy**

This study found that explicit pronunciation instruction promoted intelligibility and comprehensibility among Saudi beginning-level EFL learners, supporting a call to teach pronunciation to beginners (e.g., Zielinski & Yates, 2014). In addition, this finding demonstrates the importance of aligning pronunciation instruction to the local context. Levis (2005) similarly pointed out that “the context of the instruction directly affects how

pronunciation should be addressed” (p. 376). The following subsections focus on the Saudi context.

### **7.2.1 How Can Segmental and Suprasegmental Pronunciation Be Taught in the Saudi Beginning-Level EFL Classroom?**

The results yielded a number of practical and pedagogical implications. First, L2 pronunciation instruction promoted the intelligibility and comprehensibility of Saudi EFL learners’ speech. Therefore, policymakers and educators should focus on fostering intelligible and comprehensible pronunciation instead of native-like speech, which is not a realistic goal (Levis, 2018).

Second, not all aspects of pronunciation are important for beginning-level EFL learners, as some do not necessarily have a negative impact on intelligibility and comprehensibility (Levis, 2018). Thus, students could benefit from a needs analysis prior to pronunciation instruction (Derwing & Munro, 2015). Such an analysis would determine which segmental and suprasegmental features are most problematic and how likely they are to lessen intelligibility and comprehensibility. Once this analysis is complete, setting priorities is crucial as the time for pronunciation instruction is limited (Levis, 2018). Teachers, policymakers, and curriculum designers in the Saudi context should prioritise segmental and suprasegmental features that are important for beginners. To this end, they could follow the instructional guidelines proposed by Levis (2018, p. 186). However, some errors are more common or more difficult and should thus be prioritised. EFL teachers can follow this plan to organise their materials, prioritise vocabulary, and create opportunities to practise sounds and suprasegments using that vocabulary. They should likewise know which segmentals and suprasegmentals are most adequate to the beginning level.

Third, this study showed the effectiveness of the systematic approach proposed by Zielinski and Yates (2014). Therefore, the study encourages policymakers, curriculum designers, and EFL teachers in the Saudi context to adopt this approach by focusing on a sequence of developmental stages (listening and awareness, control, practice, and extension) to enhance beginning-level learners' intelligibility and comprehensibility.

EFL teachers may begin this process by increasing learners' awareness of the difference between /b/ and /p/ in English, as /p/ is absent from the Arabic phonemic inventory. Teachers may also make learners aware of the tendencies in lexical stress placement in English and how it is largely unpredictable, unlike Arabic. Once learners are aware of these and other basic concepts, they can move on to the control stage, which addresses physical production. For instance, teachers can work with learners to physically control the segment /v/ in a certain environment, such as the beginning of a word. Teachers then can move to the third stage to develop learners' ability to practise the target segmental and suprasegmental features in a variety of structured contexts, moving from controlled to uncontrolled ones. For instance, learners could practise intonation patterns in less-controlled structured contexts. The final stage is extension, where learners extend the pronunciation skills they have learned to other contexts, such as pronouncing features spontaneously.

Due to beginning learners' low proficiency, it could be challenging for them to perceive and produce English segmental and suprasegmental features. Therefore, it would be optimal to use various modalities to illustrate these features. Brinton (2018a) pointed out that "since pronunciation teaching involves auditory, visual and kinesthetic modalities, best practice mandates that it be taught through a multimodal approach" (p. 453). Teachers may start with auditory and visual materials to increase learners' awareness of the English phonemic inventory.

For instance, teachers may demonstrate how to correctly place stress by visually highlighting the stressed syllable in the target word (e.g., the noun “**project**” vs. the verb “**project**”).

Fourth, feedback is vital at this early stage (Zielinski & Yates, 2014), as beginning-level learners need to know when and how their pronunciation is unintelligible or incomprehensible to improve (see Darcy, 2018, p. 28). Therefore, EFL teachers should provide explicit feedback in line with Saudi EFL learners’ needs. In practical terms, teachers could draw learners’ attention to their phonetic errors (e.g., producing [b] instead [p] in “path”) and then provide explicit feedback to correct the errors. Another concrete example about the necessity of explicit feedback is that EFL teachers could explicitly explain the difference between English lexical words distinguished by stress placement (e.g., as a noun “CONduct” and as a verb “conDUCT”).

Fifth, segmental and suprasegmental features should be taught together in the Saudi beginning-level EFL classroom as they are interrelated and support one another (e.g., Lee et al., 2015; Zielinski, 2015). Teachers should know how segmentals and suprasegmentals interact and increase learners’ awareness of the importance of both in order to process L2 speech and attain intelligible and comprehensible pronunciation. For instance, the word “**produce**” might be problematic for the Saudi beginner on the segmental and prosodic levels, as they might mispronounce /p/ and misplace the stress.

Sixth, choosing the right pronunciation activities is crucial, in line with the recommendations of Zimmerman (2018). Zielinski and Yates (2014, pp. 74–76) discuss a variety of useful activities for segmentals and suprasegmentals. For example, EFL teachers can use a matching activity for lexical stress, in which learners match stress patterns to words they hear (e.g., whether stress is placed on the first or second syllable in “project”). Then, teachers can visually match stress patterns or clap for each syllable, with louder claps on the stressed syllable. Another activity is stepping out; teachers could ask learners to decide which words in

a sentence are stressed, and learners would step forward for each stressed syllable (e.g., “What **time** are they **coming**?”). Several segmental activities are likewise useful. For instance, teachers can assign a “sound of the week” to practise one problematic sound, such as /p/. Teachers could also tell a story with /p/ appearing in many words (e.g., *lap, cap, nap*). Learners would identify the sound and work in small groups to come up with their own sentences containing it.

### **7.2.2 How Can Pronunciation Gains Be Attained over Time in the Saudi Beginning-Level EFL Classroom?**

The delayed posttest showed that segmental and suprasegmental pronunciation gains after pronunciation instruction may be maintained over a long time among Saudi beginning-level EFL learners. A number of implications from these results can be drawn.

First, the duration of pronunciation instruction should be longer (e.g., one semester). However, there are several challenges to this goal. For instance, the time allocated for teaching English is limited and teachers have to cover a variety of language skills (Derwing & Munro, 2015).

Second, as the participants were beginners, it is important to continue teaching them pronunciation in later stages of English to make sure their pronunciation continues improving over the long term.

Third, as learners progress to a more advanced level, pronunciation materials should become more advanced as well. As Delicado Cantero et al. (2019) suggested, “articulatory information in isolation is not going to achieve long-lasting results without further practice” (p. 313). Therefore, Saudi EFL learners should be taught to move from a more controlled context, in which they learn to pronounce a wide range of segmental and suprasegmental features, to a less controlled context until they can produce intelligible and comprehensible speech spontaneously.

Fourth, Saudi EFL learners should explore their gains outside the classroom, as the ultimate goal of pronunciation instruction is for them to attain intelligible and comprehensible speech with interlocutors in different contexts (Levis, 2018). Since these learners are in a country where English is not necessary to communicate in everyday life, teachers and the learning material itself should encourage learners to communicate in English outside the classroom.

### **7.2.3 How Can Motivation Be Addressed in the Saudi Beginning-Level EFL Classroom?**

Intelligible and comprehensible pronunciation was linked to several factors, some beyond the type and content of instruction. One key positive factor was motivation. An important consideration, however, is that L2 learners have different motivations to improve their pronunciation (e.g., Grant, 2014a). Thus, EFL instructors, policymakers, and curriculum designers need to take these motives into account to foster Saudi learners' extrinsic and intrinsic motivation.

Three pedagogical implications can be drawn from the results. The first is based on the motivational profiles of "intended efforts" and "ideal L2 self". The suprasegmental group was more motivated than the segmental and control groups. In the same vein, the pronunciation of the suprasegmental group improved significantly more than the other groups on the posttest and delayed posttest. Therefore, EFL teachers may motivate L2 learners to put more effort into improving their pronunciation and to view intelligible and comprehensible speech as a goal. For instance, they can provide pronunciation materials connected to how L2 learners imagine themselves in the future in terms of their interests and professional goals. To address personal differences, students could be engaged in individual assignments related to their personal goals.

The second pedagogical implication is based on the third motivational dimension "ought-to L2 self". The control group was more motivated to learn English than the two

experimental groups but showed no significant improvement on the posttest and delayed posttest. This means EFL teachers should attempt to reduce pressure and fear of failure among Saudi beginner learners through a variety of strategies and activities.

The third pedagogical implication is based on the fourth motivational dimension “attitudes toward learning English”. The suprasegmental group held more positive attitudes toward learning English than the segmental and control groups. The influence of such positive attitudes could be seen in the suprasegmental group’s results in the experiment phase, as it improved significantly more than the other groups. Thus, EFL teachers should work to ensure Saudi beginning-level learners have positive attitudes about the scope, stages, and aims of pronunciation instruction, materials, activities, and instructors. In addition, policymakers and teachers are also responsible for providing a proper learning environment to help Saudi beginning-level EFL learners have positive attitudes toward learning English.

#### **7.2.4 What Do Learner Perceptions Suggest about the Effectiveness of L2 Pronunciation Instruction in the Saudi Beginning-Level EFL Classroom?**

Another pedagogical implication stemmed from learner perceptions about the effectiveness of pronunciation instruction. The quantitative and qualitative results revealed three points relevant to L2 pedagogy.

First, learners were satisfied with the intervention, finding it improved their intelligibility and comprehensibility. This could show academics, instructors, policymakers, and curriculum designers that pronunciation instruction can be effectively implemented in the Saudi beginning-level EFL classroom.

Second, learners found group practices, activities, materials, and experienced native English-speaking teachers to be the most beneficial factors in the intervention. As a result, instructors, policymakers, and curriculum designers in Saudi Arabia should consider a variety

of factors to make teaching pronunciation more effective. For instance, since teacher education is a key factor regardless of experience level, pre- and in-service teacher preparation is key for efficient learning.

Third, learners were dissatisfied with the duration of the intervention and the lack of individual practice. As noted previously, time is a challenging factor in teaching pronunciation (Darcy, 2018). Therefore, more time should be allocated to this area, when possible, with more individualised practice. Following to Darcy's (2018) suggestion, as one of possible solutions for time obstacle, EFL teachers in the Saudi beginning-level classroom can integrate pronunciation into every lesson. As an example, "a grammar unit on asking questions with *wh*-words could include specific and explicit microlessons on one or two features, such as the difference in intonation between *wh*-questions and *yes/no* questions" (p. 18). Teachers can reinforce these smaller lessons at other times during the main lesson by coming back to the target concept.

Some obstacles go beyond teacher training and preparation, such as those listed by Darcy (2018): time constraints (e.g., when and why pronunciation is taught), methodological considerations (e.g., how best to teach pronunciation features), and focus (e.g., what pronunciation features are pedagogical priorities). Although these issues appear to be institutional in nature, they should be kept in mind when designing and implementing the curriculum.

### **7.3 Implications for L2 Pronunciation Research**

While this study has shown that explicit instruction can significantly improve the pronunciation of lower-level EFL learners, researchers have been reluctant to investigate issues related to teaching pronunciation in the beginning-level EFL/ESL classroom (Zielinski & Yates, 2014). The present study encourages further empirical investigation in four areas.

First, the instruction in this study was relatively short (only three weeks), but the results were encouraging. Therefore, researchers should allocate more time, such as an entire semester, to instruction. Such studies would give a clearer picture of how well instruction produces gains over time, the ultimate goal of L2 pronunciation instruction (Derwing & Munro 2015; Thomson & Derwing, 2015).

Second, the segmental and suprasegmental results showed significant improvements in the immediate posttest, although the segment /f/ did not improve. Thus, L2 researchers could focus on both segmentals and suprasegmentals when implementing an intervention as the two are interrelated and influence each other (Levis, 2018; Zielinski, 2015) and because pronunciation instruction has been shown to be effective in both cases. In addition, the type of pronunciation materials and activities used for segmental and suprasegmental features should be a more central interest of L2 research as this issue has received less attention (Zimmerman, 2018).

Third, motivation has been significantly related to the development of L2 pronunciation (Nagle, 2018), as supported by the present study; hence, more research is required to explore other factors, as well as a wider range of motivational profiles across different stages of learning in different semesters. For instance, in the present study, the suprasegmental group was more motivated than the segmental and control groups in three motivational profiles: intended effort, ideal L2 self, and attitudes toward learning English. However, it was unclear whether participants in the suprasegmental group would be more motivated than the other groups if they were asked in later semesters.

Fourth, researchers can obtain significant quantitative and qualitative findings about the effectiveness of L2 pronunciation instruction from learners and raters' perspectives, as the results of the present study found. However, researchers should also explore learners and raters'

perceptions through interviews, as it has been outside the scope of previous research in the Saudi beginning-level EFL classroom (Al-Seghayer, 2019).

#### **7.4 Chapter Summary**

This chapter presented a number of pedagogical and research implications for incorporating pronunciation instruction into the Saudi beginning-level EFL classroom. The ability to make segmental and suprasegmental pronunciation more intelligible and comprehensible was a core implication. One of the ultimate goals of L2 pronunciation instruction is to help learners retain their gains over time. In this sense, the study found motivation was significantly related to greater L2 speech development. The implications indicated that L2 learners should no longer be a silent part of pronunciation instruction, as their input can enhance its effectiveness. However, there remain gaps in the literature on a range of issues related to L2 pronunciation instruction for beginning-level learners. The next chapter concludes the study with the study's contributions, limitations, and recommendations for future research.

## **Chapter 8: Conclusion**

### **8.1 Overview of the Study**

This study examined the effectiveness of pronunciation instruction at improving the intelligibility and comprehensibility of beginning-level Saudi EFL learners by targeting segmental and suprasegmental features. The study recruited 45 adult participants to complete a battery of pre-, immediate post-, and delayed posttests. Participants were divided into three groups. The first experimental group (the segmental group) of 15 participants received explicit pronunciation instruction on segmental features. The second experimental group (the suprasegmental group) of 15 participants received explicit instruction on suprasegmental features. The control group of 15 participants was given non-explicit pronunciation instruction to test whether explicit instruction was related to improvements in speech intelligibility and comprehensibility in the experimental groups.

The study also examined whether L2 motivation was related to the development of intelligible and comprehensible speech among the participants. In addition, they were asked about their satisfaction with the treatment and its outcomes. Finally, the study had native English-speaking raters assess participants' speech intelligibility and comprehensibility to better understand the importance of segmental and suprasegmental features.

### **8.2 Original Contributions of the Study**

This study offers a number of contributions to the field of L2 pronunciation teaching and research with implications for the Saudi beginning-level EFL classroom. First, the findings indicated that EFL pronunciation instruction is practical at the beginning stage. Although the intervention only lasted three weeks, its implementation remained effective and helped beginners make their pronunciation more intelligible and comprehensible in terms of some segmental and suprasegmental features, in particular with the suprasegmental group.

This study's focus was in line with trends in recent studies that have shifted from emphasising nativelikeness to intelligibility and comprehensibility as the goal of pronunciation instruction. Although it is unclear whether pronunciation instruction must focus on segmental or suprasegmental features, this study found both were important and should be taught because they can impede intelligibility and comprehensibility. In addition, both types of features can be improved at the same time in an integrated, interactive system. This was in line with recent calls in the L2 pronunciation literature (e.g., Zielinski, 2015).

The findings of the suprasegmental group demonstrated the ability of beginning-level EFL learners to retain pronunciation gains over time despite obstacles, such as a limited period of instruction. While the segmental group failed to retain pronunciation gains in the delayed posttest, with the exception of /v/ for intelligibility, this finding should not discourage work on long-term improvement, as the suprasegmental group succeeded in keeping improvements over time, except for the results of lexical stress for comprehensibility and rhythm for intelligibility.

Previous research has put greater emphasis on the influence of individual differences in developing L2 pronunciation (Saito et al., 2017). This study contributes to the understanding that L2 motivation is a significant predictor of the relationship between type of motivation and the development of intelligibility and comprehensibility. This was the case not only by looking at L2 motivation in a broader sense but by examining different aspects of motivational profiles (e.g., intended effort, ideal L2 self, ought-to L2 self, and attitudes toward learning English). These findings could be used to foster L2 motivation in ways that promote better pronunciation.

Participants' feedback about the intervention could help determine what challenges stand in the way of improvement and how to make the intervention more effective. Along similar lines, few studies have explored native English speakers' perceptions about L2 speakers' intelligibility and comprehensibility (e.g., Gordon & Darcy, 2016). Thus, the present study

helps address this gap. For instance, the native English-speaking raters gave valuable insights into what segmental and suprasegmental errors might impede intelligible and comprehensible communication.

### **8.3 Limitations and Recommendations for Future Research**

A number of limitations to this study could serve as avenues for future research. First of all, the pronunciation intervention was limited to three weeks. Although it helped make learners' pronunciation more intelligible and comprehensible, future work should try to teach English pronunciation throughout an entire semester. In addition, the study only examined beginner English proficiency; pronunciation instruction should continue into later stages of learning to ensure participants retain and further their gains.

This study focused only on one type of pronunciation instruction (explicit) along with certain activities and one type of corrective feedback (explicit). Although the instruction and feedback showed improvement, future research could examine the effects of different types of pronunciation instruction (e.g., implicit and explicit), different activities, and different types of corrective feedback (e.g., implicit and explicit). Although the different modalities and activities used in this study were beneficial, more advanced online platforms and resources could be used to teach English pronunciation to Saudi EFL learners in future research, as recent technological advances have been found highly useful (Brinton, 2018).

Although other language skills (e.g., writing and reading) appear related to L2 speech performance, they were beyond the scope of this study. As a result, it would be a promising direction for future research to include other language skills in pronunciation instruction.

The sample was limited to 45 participants, who were demographically homogenous. As a result, the statistical modelling must be interpreted with caution. Future work could employ a

larger sample with differences in participants' socio-economic and demographic backgrounds, as this might yield different results from the present study.

All participants were native speakers of the Sakaka dialect of Arabic found in northern Saudi Arabia. For that reason, it would be difficult to generalise the results to other Saudi provinces or beyond the Saudi EFL context. Moreover, participants were not randomly assigned because the study was carried out in intact classes. Female participants were also excluded as a consequence of logistical difficulties. Future research should consider these issues to provide further insight.

Although this study had some open-ended questions, a major component of qualitative research was not included. Therefore, future work might employ a mixed-methods design by adopting a quantitative phase (e.g., quasi-experimental approach) and a complementary qualitative phase (e.g., semi-structured interviews) to obtain a deeper understanding of the effectiveness of explicit pronunciation instruction. Nevertheless, the design adopted for this study could be applied to beginning-level learners in other EFL contexts in Saudi Arabia (e.g., outside the PYP) or outside Saudi Arabia to see whether similar results are found.

It was impossible to recruit only one person to teach pronunciation to all three groups within three weeks, as this could cause fatigue and having the same teacher over the different types of pronunciation instruction (explicit and non-explicit) could be a problem if the instructor accidentally used the wrong form of instruction with a group. However, the recruitment of two teachers was a limitation. With respect to perceptual assessment for the raters, using the same order of blocks was another limitation. However, the order of the pretest and immediate and delayed posttest audio-recordings in each block was randomised.

Segmental and suprasegmental stimulus items were limited by participants being beginning-level learners with limited access to English grammar and vocabulary. Moreover,

the scope of these features was limited to four segmental and three suprasegmental features. Other segmental and suprasegmental features that might impede intelligibility and comprehensibility could be investigated in further research. For instance, Saudi EFL learners across proficiency levels could encounter difficulties producing intelligible, comprehensible English vowels due to differences between the Arabic and English vowel systems.

However, the number of segmental and suprasegmental stimulus items was limited because it was impossible to select many stimuli that included the same characteristics (e.g., most frequent monosyllabic real words and real full sentences with a simple grammatical structure containing the target segmental and suprasegmental features) while closely aligning with participants' basic language proficiency. Although it has been debated whether using different stimuli in each test would be better (Colantoni et al., 2015), another limitation was using the same segmental and suprasegmental stimuli across the tests. As a result, future research might employ more stimulus items with more disyllabic and trisyllabic words that are less frequent and have a more complex grammar and structure. Different stimuli could also be used in each test.

The types of tasks used to measure intelligibility and comprehensibility represented another limitation; the tasks were well-controlled because participants were beginning-level learners. For instance, two-way forced choice identification of minimal pairs (heard as /p/ or /b/) was used to measure intelligibility, while rating sentences read aloud (from 1 "very difficult to understand" to 9 "very easy to understand") was used for comprehensibility. Although these tasks elicited rich data, future research could use less-controlled tasks for measuring intelligibility, in particular with intermediate and advanced learners, such as a transcription task (which was used for measuring rhythmic patterns in this study). With respect to measuring

comprehensibility, the Likert scale must be clearly defined. However, this issue has been controversial in L2 pronunciation assessment (e.g., Thomson, 2018).

Future work could also be based on an acoustical analysis, which was outside the scope of this study. Moreover, as this study only examined the speech production of beginning-level learners, future studies could explore whether there is a link between students' speech production and perception abilities. The importance of this relationship stems from the nature of the mechanisms of these two modalities that interact with each other at least on the underlying level (Flege & Bohn, 2021). Nagle and Baese-Berk (2021) also called for linking the perception and production modalities in L2 speech learning research and taking into account, for instance, L1 and L2 contrasts and the different phonetic cues between the L1 and L2. In addition, a diagnostic test of English pronunciation, an individualised test, or a needs analysis could offer an excellent benchmark to help teachers and students more precisely identify student pronunciation problems prior to implementing pronunciation instruction.

Finally, it is worth mentioning that L2 motivation in the present study was assessed domain-generally by adopting Taguchi et al.'s (2009) motivation questionnaire to explore four motivational profiles of Saudi beginning-level EFL learners. Thus, a promising direction for future research would be to adopt a motivation survey developed based on a domain-specific approach. Recent motivation research has suggested that motivation can be explored domain-specifically, as students could have different motivational orientations in each area of English skills, including pronunciation (e.g., Baran-Łucarz, 2016, 2017). This suggestion should be considered in future research, as investigating motivation profiles through a domain-specific approach could deepen our understanding of the relationship between L2 motivation and oral development.

## References

- Abercrombie, D. (1949). Teaching pronunciation. *English Language Teaching*, 3, 113–122.
- Abu-Ghararah, B. (2014). Problems with speaking activities in the Saudi EFL classroom. *Arab World English Journal*, 5(4), 276–287.
- Ahmad, A., Karim, S., & Khokhar, M. I. (2016). Attitudes and perceptions of the Saudi EFL learners about the Pakistani EFL teachers. *Journal of Educational Research*, 19(2), 62–77.
- Ahmad, J., & Muhiburrahman, M. (2013). Teachers' perspectives on errors in English consonant sounds by Saudi EFL learners. *Asian Journal of Humanities and Social Sciences*, 1(3), 150–163.
- Ahmad, J., & Nazim, M. (2014). Saudi EFL learners' views on English pronunciation. *Asian Journal of Studies and Social Studies*, 2(2), 355–366.
- Al-Ahdal, H., Mohammed, A. A., Al-Hattami, A. A., Al-Awaid, S. A. A., & Al-Mashaqba, N. J. A. H. (2015). Pronunciation for the Arab learners of EFL: Planning for better outcomes. *English Language Teaching*, 8(10), 100–106.
- Alameen, G. (2014). *The effectiveness of linking instruction on NNS speech perception and production* [Unpublished doctoral dissertation]. Iowa State University, Ames, Iowa.
- Alamri, M. (2011). Higher education in Saudi Arabia. *Journal of Higher Education Theory and Practice*, 11(4), 88–91.
- Alasmari, A. R. (2013). Saudi university undergraduates' language learning attitudes: A preparatory year perspective. *International Journal of Asian Social Science*, 3(11), 2288–2306.
- Alfallaj, F. (2013). The use of Linguistics in teaching pronunciation to Saudi students. *Journal of Humanities*, 14(2), 134–143.
- Algethami, G. (2016). Long term effect of phonetic instruction on the production of /p/ by EFL Arab learners: An exploratory study. *Arab World English Journal*, 7(4), 297–304.
- Algethami, G. (2017). Effect of explicit pronunciation instruction on the degree of perceived foreign accent in the speech of EFL learners. *Research in Language*, 15(3), 253–263.
- Alghamdi, O. S. (2015). Satisfaction of preparatory year students with university services. *World Journal of Education*, 5(5), 117–129.
- Al-Hajailan, T. (2006). *Teaching English in Saudi Arabia*. Aldar Alsawlatiah.
- Al-Harbi, A. (2006). *Training needs of English teachers in Al-Qassim province* [Unpublished master's thesis]. King Saud University.
- Alharbi, E. A. R. (2016). Higher education in Saudi Arabia: Challenges to achieving world-class recognition. *International Journal of Culture and History*, 2(4), 169–172.
- Alharbi, H. A. (2015). Improving students' English speaking proficiency in Saudi public schools. *International Journal of Instruction*, 8(1), 105–116.
- Alharbi, T. N. (2014). *The perception of English syllable-final nasals by Saudi ESL learners* [Unpublished master's thesis]. Southern Illinois University.
- Al-Hazmi, S. (2003). EFL teacher preparation program in Saudi Arabia: Trend and challenges. *TESOL Quarterly*, 37(2), 341–344.
- Ali, E. M. T. (2015). The elimination of pronunciation problems of English vowels of Saudi students of English resulting from complex letter-sound relationship. *International Journal of Language and Literature*, 3(1), 35–41.
- Aljumah, F. (2011). Developing Saudi EFL students' oral skills: An integrative approach. *English Language Teaching*, 4(3), 84–89.

- Almbark, R. (2012). *The perception and production of SSBE vowels by Syrian Arabic learners: The foreign language model* [Unpublished doctoral dissertation]. University of York.
- Alqahtani, M. S. (2018). Teaching English in Saudi Arabia. In C. Moskovsky & M. Picard (Eds.), *English as foreign language in Saudi Arabia: New insights into teaching and learning English* (pp. 120–137). Routledge.
- Arabai, F. (2016). Factors underlying low achievement of Saudi EFL learners. *International Journal of English Linguistics*, 6(3), 21–37.
- Alrasheed, A. (2010). *Teaching skills for newly qualified teachers*. Al-Rushed Press.
- Alrashidi, O., & Phan, H. (2015). Education context and English teaching and learning in the Kingdom of Saudi Arabia: An overview. *English Language Teaching*, 8(5), 33–44.
- Al-Seghayer, K. (2011). *English teaching in Saudi Arabia: Status, issues, and challenges*. Hala Printing Co.
- Al-Seghayer, K. (2014). The four most common constraints affecting English teaching in Saudi Arabia. *International Journal of English Linguistics*, 4(5), 17–26.
- Al-Seghayer, K. (2015). Salient key features of actual English instructional practices in Saudi Arabia. *English Language Teaching*, 8(6), 89–99.
- Al-Seghayer, K. (2019). **Unique challenges Saudi EFL learners face**. *Studies in English Language Teaching*, 7(4), 490–515. <https://doi.org/10.22158/selt.v7n4p490>
- Alshahrani, M. (2016). A brief historical perspective of English in Saudi Arabia. *Journal of Literature, Languages and Linguistics*, 26, 43–47.
- Al-Shehri, A. S. (2009). Motivation and vision: The relation between the ideal L2 self, imagination and visual style. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 164–171). Multilingual Matters.
- Alsofyani, M., & Algethami, G. (2017). Exploring EFL teachers' beliefs and practices regarding pronunciation teaching in a Saudi setting. *Arab World English Journal*, 8(4), 384–400.
- Altaha, F. M. (1995). Pronunciation errors made by Saudi university students learning English. *International Journal of Applied Linguistics*, 109(1), 110–123.
- Altalib, A. (2019). L2 motivation in ESP and EGP courses: An investigation of L2 motivational selves among learners of English in Saudi Arabia. *Australian Journal of Applied Linguistics*, 2(1), 1–16. <https://dx.doi.org/10.29140/ajal.v2n1.113>
- Althobaiti, H. (2017). How can in-service development take a part in Saudi English teacher development? *British Journal of Education*, 5(3), 21–29.
- Al-Wossabi, S. (2016). SLA classroom research and EFL teaching practices of oral skills. *Theory and Practice in Language Studies*, 6(11), 2061–2067.
- Alyami, R. H. (2014). Educational reform in the Kingdom of Saudi Arabia: Tatweer schools as a unit of development. *Literacy Information and Computer Education Journal*, 5(2), 1424–1433.
- Alzinaidi, M. H., & Abdel Latif, M. M. M. (2019). Diagnosing Saudi students' English consonant pronunciation errors. *Arab World English Journal*, 10(4), 180–193. <https://dx.doi.org/10.24093/awej/vol10no4.14>
- Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. A. (2019). *Introduction to research in education* (10th ed.). Cengage Learning.
- Baran-Lucarz, M. (2016). Conceptualizing and measuring the construct of pronunciation anxiety: Results of a pilot study. In M. Pawlak (Ed.), *Classroom-oriented research* (pp. 39–56). Springer.

- Baran-Łucarz, M. (2017). FL pronunciation anxiety and motivation: Results of a mixed-method study. In E. Piechurska-Kuciel, E. Szymańska-Czaplak, & E. Szyszka (Eds.), *At the crossroads: Challenges of foreign language learning* (pp. 107–133). Springer.
- Barros, A. M. (2003). *Pronunciation difficulties in the consonant system experienced by Arabic speakers when learning English after the age of puberty* [Unpublished master's thesis]. West Virginia University.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.  
<http://doi.org/10.18637/jss.v067.i01>
- Best, C. T. (1993). Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. W. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 289–304). Kluwer Academic Publishers.
- Best, C. T. (1994a). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 8, pp. 217–304). Ablex.
- Best, C. T. (1994b). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). MIT Press.
- Best, C. T. (1995). A direct realist view of cross-language speech. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 171–203). York Press.
- Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross language perception of approximants. *Journal of Phonetics*, 20(3), 305–330.
- Best, C. T., & Tyler, M. D. (2007). Non-native and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34). John Benjamins.
- Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109, 775–794.
- Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for non-native speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14(3), 345–360.
- Binturki, T. A. (2008). *Analysis of pronunciation errors of Saudi ESL learners* [Unpublished master's thesis]. Southern Illinois University.
- Boersma, P., & Weenink, D. (2019). Praat: Doing phonetics by computer [Computer program]. Version 6.0.50. Retrieved 10 April 2019 from <http://www.praat.org/>
- Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19, 447–465.
- Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, 101(4), 2299–2310.

- Brinton, D. M. (2014). Epilogue to the myths: Best practices for teachers. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 225–242). University of Michigan Press.
- Brinton, D. M. (2018a). Innovations in pronunciation teaching. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 449–461). Routledge.
- Brinton, D. M. (2018b). Reconciling theory and practice. *CATESOL Journal*, 30(1), 283–300.
- Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(4), 155–180.
- Browman, C. P., & Goldstein, L. M. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201–251.
- Celce-Murcia, M., Brinton, D., Goodwin, J., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd ed.). Cambridge University Press.
- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1997). Direct approaches in L2 instruction: A turning point in communicative language teaching? *TESOL Quarterly*, 31(1), 141–152.
- Chun, D. M., Jiang, Y., & Ávila, N. (2013). Visualization of tone for learning Mandarin Chinese. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference* (pp. 77–89). Iowa State University.
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Colantoni, L., Steele, J., & Escudero, P. (2015). *Second language speech: Theory and practice*. Cambridge University Press.
- Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect*, 21(1), 46–66.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative and mixed methods approaches* (4th ed.). Sage.
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative and mixed methods approaches* (5th ed.). Sage.
- Crookes, G., & Schmidt, R. W. (1991). Motivation: Reopening the research agenda. *Language Learning*, 41(4), 469–512.
- Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369–382.
- Cutler, A. (2015). Lexical stress in English pronunciation. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 106–124). Wiley Blackwell.
- Daif-Allah, A. S., & Khan, M. I. (2016). The impact of open discussion sessions on enhancing the oral communicative abilities of Saudi English language majors at Buraydah Community College. *English Language Teaching*, 9(6), 108–122.
- Darcy, I. (2018). Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal*, 30(1), 13–46.
- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51–62.
- Delicado Cantero, M., Steed, W., & Herrero de Haro, A. (2019). Spanish pronunciation and teacher training: Challenges and suggestions. In R. Rao (Ed.), *Key issues in the teaching of Spanish pronunciation: From description to pedagogy* (pp. 304–323). Routledge.

- Derwing, T. M. (2003). What do ESL students say about their accents? *Canadian Modern Language Review*, 59(4), 547–567. <https://doi.org/10.3138/cmlr.59.4.547>
- 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.
- Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, 63(2), 163–185.
- Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.
- 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., Munro, M. J., Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning*, 64(3), 526–548. <https://doi.org/10.1111/lang.12053>
- Dörnyei, Z. (1994). Motivation and motivating in a foreign language. *Modern Language Journal*, 78(3), 273–284.
- Dörnyei, Z. (2005). *The psychology of the language learning: Individual differences in second language acquisition*. Lawrence Erlbaum Associates.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters.
- Dörnyei, Z. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). Routledge.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Longman.
- Ebad, R. (2014). The role and impact of English as a language and a medium of instruction in Saudi higher education institutions: Students-instructors perspective. *Studies in English Language Teaching*, 2(2), 140–148.
- Elyas, T., & Al-Grigri, W. H. (2014). Obstacles to teaching English in Saudi Arabia public schools: Teachers’ and supervisors’ perceptions. *International Journal of English Language Teaching*, 2(3), 74–89.
- Elyas, T., & Picard, M. (2010). Saudi Arabian educational history: Impacts on English language teaching. *Education, Business and Society: Contemporary Middle Eastern Issues*, 3(2), 136–145.
- Elyas, T., & Picard, M. (2018). A brief history of English in Saudi Arabia. In C. Moskovsky & M. Picard (Eds.), *English as foreign language in Saudi Arabia: New insights into teaching and learning English* (pp. 70–84). Routledge.
- Escudero, P. (2007). Second-language phonology: The role of perception. In M. C. Pennington (Ed.), *Phonology in context* (pp. 109–134). Palgrave Macmillan.
- Field, J. (2005). Intelligibility and the listener: The role of lexical stress. *TESOL Quarterly*, 39(3), 399–423.
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, 15(1), 47–65.
- Flege, J. E. (1988). The production and perception of foreign language speech sounds. *Human Communication and Its Disorders: A Review*, 1, 224–401.

- Flege, J. E. (1995). Second language speech learning theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- Flege, J. E. (2003). Methods for assessing the perception of vowels in a second language. In E. Fava & A. Mioni (Eds.), *Issues in clinical linguistics* (pp. 19–44). Padova University Press.
- Flege, J. E., & MacKay, I. R. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26(1), 1–34.
- Flege, J. E., & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, 24(2), 125–146.
- Flege, J. E., Schirru, C., & MacKay, I. R. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40(4), 467–491.
- Flege, J., & Bohn, O. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (pp. 3–83). Cambridge University Press. <https://doi.org/10.1017/9781108886901.002>
- Foote, J. A., & Trofimovich, P. (2018). Second language pronunciation learning: An overview of theoretical perspectives. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 75–90). Routledge.
- Fowler, C. A. (1986). An event approach to the study of speech perception from a direct-realist perspective. *Journal of Phonetics*, 14, 3–28.
- Fowler, C. A. (1989). Real objects of speech perception: A commentary on Diehl and Kluender. *Ecological Psychology*, 1(2), 145–160.
- French, L. M., Gagné, N., & Collins, L. (2020). Long-term effects of intensive instruction on fluency, comprehensibility and accentedness. *Journal of Second Language Pronunciation*, 6(3), 380–401. <https://doi.org/10.1075/jslp.20026.fre>
- Frost, D. (2011). Stress and cues to relative prominence in English and French: A perceptual study. *Journal of the International Phonetic Association*, 41(1), 67–84.
- Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). *irr: Various coefficients of interrater reliability and agreement*. <https://cran.r-project.org/web/packages/irr/index.html>
- Gardner, R. C. (1985). *Social psychology and language learning: The role of attitudes and motivation*. Edward Arnold.
- General Authority for Statistics. (2019). *Statistical yearbook of 2019*. <https://www.stats.gov.sa/en/1006>
- Gibson, E. J. (1991). *An odyssey in learning and perception*. MIT Press.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton Mifflin.
- Gilbert, J. B. (2008). *Teaching pronunciation: Using the prosody pyramid*. [https://www.pronunciationinaction.com/downloads/JGilbert\\_Teaching%20Pronunciation.pdf](https://www.pronunciationinaction.com/downloads/JGilbert_Teaching%20Pronunciation.pdf)
- Gordon, J., & Darcy, I. (2016). The development of comprehensible speech in L2 learners: A classroom study on the effects of short-term pronunciation instruction. *Journal of Second Language Pronunciation*, 2, 56–92. <https://doi.org/10.1075/jslp.2.1.03gor>
- Grant, L. (2014a). Myth 5: Students would make better progress in pronunciation if they just practiced more. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 141–159). University of Michigan Press.
- Grant, L. (2014b). *Pronunciation myths: Applying second language research to classroom teaching*. University of Michigan Press.
- Grant, L., & Yu, E. E. (2016). *Well said intro: Pronunciation for clear communication* (2nd ed.). National Geographic Learning and Cengage Learning.

- Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*(5), 2711–2724.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, *38*, 201–223.
- Hamad, M. M. (2013). Factors negatively affect speaking skills at Saudi colleges for girls in the south. *English Language Teaching*, *6*(12), 87–97.
- Hameed, P. F. M., & Aslam, M. S. (2015). Pronunciation as a stumbling block for the Saudi English learners: An analysis of the problems and some remedies. *Theory and Practice in Language Studies*, *5*(8), 1578–1585.
- Hansen Edwards, J. G. (2017). Pronunciation and individual differences. In O. Kang, R. Thomson, & J. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 385–398). Routledge.
- Hardison, D. M. (2004). Generalization of computer-assisted prosody training: Quantitative and qualitative findings. *Language Learning & Technology*, *8*(1), 34–52.
- Hayes-Harb, R., & Hacking, J. F. (2015). Beyond rating data: What do listeners believe underlies their accentedness judgments? *Journal of Second Language Pronunciation*, *1*(1), 43–64. <https://doi.org/10.1075/jslp.1.1.02hay>
- Hayes-Harb, R., & Watzinger-Tharp, J. (2012). Accent, intelligibility, and the role of the listener: Perceptions of English-accented German by native German speakers. *Foreign Language Annals*, *2*(45), 260–282.
- Hayes-Harb, R., Smith, B. L., Bent, T., & Bradlow, A. R. (2008). The interlanguage speech intelligibility benefit for native speakers of Mandarin: Production and perception of English word-final voicing contrasts. *Journal of Phonetics*, *36*(4), 664–679.
- Higgins, E. T. (1987). Self-discrepancy: A theory relating self and affect. *Psychological Review*, *94*(3), 319–340. <http://dx.doi.org/10.1037/0033-295x.94.3.319>
- Hirata, Y. (2004). Computer assisted pronunciation training for native English speakers learning Japanese pitch and durational contrasts. *Computer Assisted Language Learning*, *17*(3-4), 357–376.
- Huensch, A., & Thompson, A. S. (2017). Contextualizing attitudes toward pronunciation: Foreign language learners in the United States. *Foreign Language Annals*, *50*(2), 410–432.
- Huensch, A., & Tremblay, A. (2015). Effects of perceptual phonetic training on the perception and production of second language syllable structure. *Journal of Phonetics*, *52*, 105–120.
- Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English-speaking graduate students. *The Canadian Modern Language Review*, *64*, 555–580.
- Isaacs, T. (2009). Integrating form and meaning in L2 pronunciation instruction. *TESL Canada Journal*, *27*(1), 1–12.
- Isaacs, T., & Thomson, R. I. (2013). Rater experience, rating scale length, and judgments of L2 pronunciation: Revisiting research conventions. *Language Assessment Quarterly*, *10*(2), 135–159. <https://doi.org/10.1080/15434303.2013.769545>
- Iverson, P., Hazan, V., & Bannister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English/r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America*, *118*(5), 32673278.

- Iverson, P., Pinet, M., & Evans, B. G. (2012). Auditory training for experienced and inexperienced second-language learners: Native French speakers learning English vowels. *Applied Psycholinguistics*, 33(1), 145–160.
- Jenkins, J. (2000). *The phonology of English as an international language: New models, new norms, new goals*. Oxford University Press.
- Jenkins, J. (2013). *English as a lingua franca in the international university: The politics of academic English language policy*. Routledge.
- Kang, O., & Ginther, A. (Eds.). (2018). *Assessment in second language pronunciation*. Routledge.
- Kang, O., Thomson, R. I., & Moran, M. (2018). Empirical approaches to measuring the intelligibility of different varieties of English in predicting listener comprehension. *Language Learning*, 68, 115–146. <https://doi.org/10.1111/lang.12270>
- Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., & Golestani, N. (2015). The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds a. *The Journal of the Acoustical Society of America*, 138(2), 817–832.
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2019). The language experience and proficiency questionnaire (LEAP-Q): Ten years later. *Bilingualism: Language and Cognition*.
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, comprehensibility and accentedness of L2 speech: The role of listener experience and semantic context. *Canadian Modern Language Review*, 64, 459–489.
- Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language Awareness*, 19, 171–185.
- Kenworthy, J. (1987). *Teaching English pronunciation*. Oxford University Press.
- Khan, I. (2013). Speaking skills and teaching strategies: The case of an EFL classroom. *Elixir International Journal*, 58(10), 14557–14560.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kuznetsova, A., Brockhoff, P., & Christensen, R. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. <http://doi.org/10.18637/jss.v082.i13>
- Lee, A., & Lyster, R. (2017). Can corrective feedback on second language speech perception errors affect production accuracy? *Applied Psycholinguistics*, 38(2), 371–393.
- Lee, B., Plonsky, L., & Saito, K. (2020). The effects of perception- vs. production-based pronunciation instruction. *System*, 88, 1–13. <https://doi.org/10.1016/j.system.2019.102185>
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36(3), 345–366. <https://doi.org/10.1093/applin/amu040>
- Lepage, A., & Busà, M. G. (2014). Intelligibility of English L2: The effects of incorrect word stress placement and incorrect vowel reduction in the speech of French and Italian learners of English. In *Proceedings of the International Symposium on the Acquisition of Second Language Speech Concordia Working Papers in Applied Linguistics* (pp. 387–400).
- Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly*, 39(3), 369–377. <https://doi.org/10.2307/3588485>

- Levis, J. M. (2015). Learners' views of social issues in pronunciation learning. *Journal of Academic Language and Learning*, 9(1), 42–55.
- Levis, J. M. (2016). Research into practice: How research appears in pronunciation teaching materials. *Language Teaching*, 49(3), 423–437.
- Levis, J. M. (2018). *Intelligibility, oral communication and the teaching of pronunciation*. Cambridge University Press.
- Levis, J. M. (2020a). Changes in L2 pronunciation: 25 years of intelligibility, comprehensibility, and accentedness. *Journal of Second Language Pronunciation*, 6(3), 277–282.
- Levis, J. M. (2020b). Revisiting the intelligibility and nativeness principles. *Journal of Second Language Pronunciation*, 6(3), 310–328. <https://doi.org/10.1075/jslp.20050.lev>
- Levis, J. M., & Wichmann, A. (2015). English intonation: Form and meaning. In M. Reed and J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 139–155). Wiley Blackwell.
- Lima, E. F. (2017). A review of the new *Well Said* series [Review of the book *Well said intro: Pronunciation for clear communication 2nd ed.*, by L. Grant & E. Yu]. *Journal of Second Language Pronunciation*, 3(2), 293–297.
- Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, 65, 185–207. <https://doi.org/10.1111/lang.12117>
- Liton, H. A. (2013). EFL teachers' perceptions, evaluations and expectations about English language courses as EFL in Saudi universities. *International Journal of Instruction*, 6(2), 19–34.
- Low, E. L. (2014). *Pronunciation for English as an international language: From research to practice*. Routledge.
- Low, E. L. (2015). The rhythmic patterning of English(es): Implications for pronunciation teaching. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 125–138). Wiley Blackwell.
- Lüdecke, D. (2018). SjPlot: Data visualization for statistics in social science [R package sjPlot version 2.6.2]. <https://doi.org/10.5281/zenodo.1308157>
- Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design* (2nd ed.). Routledge.
- Mahboob, A., & Elyas, T. (2014). English in the Kingdom of Saudi Arabia. *World Englishes*, 33(1), 128–142.
- Mahdi, D. A. (2015). Motivating reluctant EFL students to talk in class: Strategies and tactics. *Theory and Practice in Language Studies*, 5(8), 1703–1709.
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50, 940–967.
- Markus, H., & Nurius, P. (1986). Possible selves. *American Psychologist*, 41(9), 954–969. <http://dx.doi.org/10.1037/0003-066x.41.9.954>
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30–46. <https://doi.org/10.1037/1082-989X.1.1.30>
- Mennen, I. (2015). Beyond segments: Towards an L2 intonation learning theory (LILt). In E. Delais-Roussarie, M. Avanzi, & S. Herment (Eds.), *Prosody and language in contact: L2 acquisition, attrition and languages in multilingual situations* (pp. 171–188). Springer-Verlag. <https://doi.org/10.1007/978-3-662-45168-7>

- Ministry of Education. (2013). *English language curriculum for elementary, intermediate and secondary schools in the Kingdom of Saudi Arabia Grades 4–12, 2014–2020*. <https://eelyanbu.files.wordpress.com/2014/02/english-language-curriculum-for-schools-in-the-ksa-final.pdf>
- Ministry of Education. (2019). *Establishment*. <https://www.moe.gov.sa/en/TheMinistry/AboutMinistry/Pages/EstablishmentoftheMinistryofEducation.aspx>.
- Mitchell, B., & Alfuraih, A. (2017). English language teaching in the Kingdom of Saudi Arabia: Past, present and beyond. *Mediterranean Journal of Social Sciences*, 8(2), 317–325.
- Moskovsky, C., Racheva, S., Assulaimani, T., & Harkins, J. (2016). The L2 motivational self system and L2 achievement: A study of Saudi EFL learners. *The Modern Language Journal*, 100(3), 1–14. <https://doi.org/10.1111/modl.12340>
- Moyer, A. (1999). Ultimate attainment in L2 phonology: The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21(1), 81–108.
- Moyer, A. (2014). Exceptional outcomes in L2 phonology: The critical factors of learner engagement and self-regulation. *Applied Linguistics*, 35(4), 418–440.
- Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, 36(1), 39–66.
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Special reprint of the decade's best papers in Language Learning*, 49, 285–310.
- Munro, M. J., & Derwing, T. M. (2015). Intelligibility in research and practice: Teaching priorities. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 377–396). Wiley-Blackwell Publishers.
- Munro, M. J., & Derwing, T. M. (2020). Foreign accent, comprehensibility and intelligibility, redux. *Journal of Second Language Pronunciation*, 6(3), 283–309. <https://doi.org/10.1075/jslp.20038.mun>
- Munro, M. J., & T. M. Derwing. (2006). The functional load principle in ESL pronunciation instruction: An exploratory study. *System* 34, 520–531.
- Munro, M. J., Derwing, T. M., & Morton, S. L. (2006). The mutual intelligibility of L2 speech. *Studies in Second Language Acquisition*, 28(1), 111–131.
- Murphy, J. (2017). Suprasegmentals: Thought groups, prominence, word stress, intonation and pitch jumps. In J. Murphy (Ed.), *Teaching the pronunciation of English* (pp. 31–69). University of Michigan Press.
- Murphy, J., & 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.
- Nagle, C. (2018). Motivation, comprehensibility, and accentedness in L2 Spanish: Investigating motivation as a time-varying predictor of pronunciation development. *The Modern Language Journal*, 102(1), 199–217. <https://doi.org/10.1111/modl.12461>
- Nagle, C., & Baese-Berk, M. (2021). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*. <https://doi.org/10.1017/S0272263121000371>
- O'Brien, M. G. (2021). Ease and difficulty in L2 pronunciation teaching: A mini-review. *Frontiers in Communication*, 5, 1–7. <https://doi.org/10.3389/fcomm.2020.626985>

- Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26–45.
- Oxford, R., & Shearin, J. (1994). Language learning motivation: Expanding the theoretical framework. *The Modern Language Journal*, 78(1), 12-28.  
<https://doi.org/10.2307/329249>
- Pennington, M. C. (2021). Teaching pronunciation: The state of the art 2021. *RELC Journal*, 52(1), 3–21. <https://doi.org/10.1177/00336882211002283>
- Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics*, 6, 179–206.
- Piske, T., MacKay, I. R., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191–215.
- Prator, C. H., & Robinett, B. J. (1985). *A manual of American English pronunciation* (4th ed.). Holt, Rinehart, & Winston.
- Qualtrics. (2020). [Software]. <http://www.qualtrics.com>
- R Core Team. (2019). *R: A language and environment for statistical computing*.  
<http://www.R-project.org/>
- Rahman, M. M., & Alhaisoni, E. (2013). Teaching English in Saudi Arabia: Prospects and challenges. *Academic Research International*, 4(1), 112–118.
- Rajab, H. (2013). Developing speaking and writing skills of L1 Arabic EFL learners through teaching of IPA phonetic codes. *Theory and Practice in Language Studies*, 3(4), 653–659.
- Reed, J., & Levis, J. (Eds.). (2015). *The handbook of English pronunciation*. Wiley-Blackwell Publishers.
- Rogerson-Revell, P. (2018). English vowels and consonants. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 93–121). Routledge.
- Ruellot, V. (2011). Computer-assisted pronunciation learning of French /u/ and /y/ at the intermediate level. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference* (pp. 199–213). Iowa State University.
- Ryan, S. (2006). Language learning motivation within the context of globalization: An L2 self within an imagined global community. *Critical Inquiry in Language Studies: An International Journal*, 3(1), 23–45. [https://doi.org/10.1207/s15427595cils0301\\_2](https://doi.org/10.1207/s15427595cils0301_2)
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *TESOL Quarterly*, 46, 842–854.  
<https://doi.org/10.1002/tesq.67>
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners of English. *Language Learning*, 62(2), 595-633.
- Saito, K., Dewaele, J.-M., & Hanzawa, K. (2017). A longitudinal investigation of the relationship between motivation and late second language speech learning in classroom settings. *Language and Speech*, 60(4), 614–632.  
<https://doi.org/10.1177/0023830916687793>
- Saito, K., Dewaele, J.-M., Abe, M. and In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: A cross-sectional and longitudinal study. *Language Learning*, 68(3), 709–743. <https://doi.org/10.1111/lang.12297>

- Saito, K., Trofimovich, P., & Isaacs, T. (2016). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. *Applied Psycholinguistics*, 37(2), 217–240. <https://doi.org/10.1017/S0142716414000502>
- Scales, J., Wennerstrom, A., Richard, D., & Wu, S. (2006). Language learners' perceptions of accent. *TESOL Quarterly*, 40(4), 715–738. <https://doi.org/10.2307/40264305>
- Sedarous, Y., & Namboodiripad, S. (2020). Using audio stimuli in acceptability judgment experiments. *Language and Linguistics Compass*, 14(8). <https://doi.org/10.1111/lnc3.12377>
- Setter, J., & Sebina, B. (2018). English lexical stress, prominence and rhythm. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 137–153). Routledge.
- Shah, S. R., & Al-Bargi, A. (2016). Intelligibility: The goal of language learning and teaching. *Journal for the Study of English Linguistics*, 4(1), 1–18.
- Sheppard, B. E., Elliott, N. C., & Baese-Berk, M. M. (2017). Comprehensibility and intelligibility of international student speech: Comparing perceptions of university EAP instructors and content faculty. *Journal of English for Academic Purposes*, 26, 42–51.
- Springall, J. (2002). *Pronunciation project* [Unpublished manuscript].
- Swan, M., & Smith, B. (2001). *Learner English: A teacher's guide to interference and other problems* (2nd ed.). Cambridge University Press.
- Taguchi, T., Magid, M., & Papi, M. (2009). The L2 motivational self system amongst Chinese and Japanese learners of English: A comparative study. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 66–97). Multilingual Matters.
- Thomson, R. (2018). Measurement of accentedness, intelligibility and comprehensibility. In O. Kang & A. Ginther (Eds.), *Assessment in second language pronunciation* (pp. 11–29). Routledge.
- Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Calico Journal*, 28(3), 744–765.
- Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36(3), 326–344. <https://doi.org/10.1093/applin/amu076>
- Tremblay, A. (2009). Phonetic variability and the variable perception of L2 word stress by French Canadian listeners. *International Journal of Bilingualism*, 13(1), 35–62.
- Trofimovich, P., Kennedy, S., & Foote, J. A. (2015). Variables affecting L2 pronunciation development. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 353–373). Wiley-Blackwell.
- Ullah, F. (2017). Socio-cultural constraints in learning English language at Jazan University, Kingdom of Saudi Arabia. *International Journal of Language and Linguistics*, 5(2), 29–33. <https://doi.org/10.11648/j.ijll.20170502.11>
- Ur Rahman, M. M., & Alhaisoni, E. (2013). Teaching English in Saudi Arabia: prospects and challenges. *Academic Research International*, 4(1), 112–118.
- Ushioda, E. (2006). Language motivation in a reconfigured Europe: Access, identity, autonomy. *Journal of Multilingual and Multicultural Development*, 27(2), 148–161.
- Vo, S. C., Vo, Y. T. H., & Vo, Q. T. (2014). The amount of English use: Effects on L2 speech. *TESL-EJ*, 17(4), 1–18.

- Wang, X. (2020). Segmental versus suprasegmental: Which one is more important to teach? *RELC Journal*, 1–9. <https://doi.org/10.1177/0033688220925926>
- Wennerstrom, A. (2018). Intonation and language learning. In O. Kang, R. I. Thomson, & J. M. Murphy (Eds.), *The Routledge handbook of contemporary English pronunciation* (pp. 154–168). Routledge.
- Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the Internet: A tutorial review. *The Journal of Life and Environmental Sciences*, 3. <https://doi.org/10.7717/peerj.1058>
- Word Frequency. (n.d.). *Word frequency data*. <https://www.wordfrequency.info/intro.asp>
- Zhang, R., & Yuan, Z. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition*, 42(4), 905–918. <https://doi.org/10.1017/S0272263120000121>
- Zielinski, B. W. (2008). The listener: No longer the silent partner in reduced intelligibility. *System*, 36, 69–84. <https://doi.org/10.1016/j.system.2007.11.004>
- Zielinski, B. W. (2015). The segmental/suprasegmental debate. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 397–412). Wiley-Blackwell.
- Zielinski, B., & Yates, L. (2014). Myth 2: Pronunciation instruction is not appropriate for beginning level learners. In L. Grant (Ed.), *Pronunciation myths: Applying second language research to classroom teaching* (pp. 56–79). Michigan University Press.
- Zimmerman, E. (2018). Reconsidering assumptions of beginner teachers' needs: An examination of commonly used pronunciation textbooks. *CATESOL Journal*, 30(1), 231–248.



- **Language Background**

1. What is your first language and your native dialect?

---



---

2. How many languages have you studied/been exposed to? Please list them.

---



---

3. At what age did you start studying English? What was the reason?

---



---

4. Have you ever studied English prior to enrolling in the English program at the preparatory year? If yes, please fill out the following table:

Yes

No

Location of study (e.g., high school, private institute, online)	Length of study (e.g., one semester)	Type of course (e.g., an introduction to general English skills)

5. Have you ever been exposed to English pronunciation instruction or phonetics and phonology training? If yes, please fill out the following table:

Yes

No

Pronunciation feature (vowels/ consonants/ stress/ rhythm/ intonation/ other)	Length of instruction (e.g., one week)	Location of instruction (e.g., classroom/ lab/ online)

6. Have you ever had native English-speaking teachers while studying English? If yes, please fill out the following table:

Yes

No

Nationality	Which course	Which studying/learning stage (e.g., high school)

7. Have you ever been in an English-speaking country? If yes, please fill out the following table:

Yes

No

Country	Length of trip	Reason (tourism/visit friends or relatives/medicine, etc.)


8. Do you use English outside of class? If yes, please fill out the following table:

Yes

No

How many hours per week	Using English via (websites/ apps/ TV, etc.)	Reason (study/ pleasure/ watching movies or news, etc.)

9. Do you face difficulties to be understood by your native English-speaking teachers? If yes, please explain these difficulties from your own perspective.

Yes

No

---



---

## Appendix B: L2 Motivation Questionnaire (English Version)

Dear Participant,

I would like to thank you for agreeing to participate in the present study. The primary purpose of this questionnaire is to explore how much you are motivated to learn English. I would also like to know what you feel when you learn English. I would like to bring to your attention that this study is not meant to assess your English proficiency or serve any other type of examination as your information will only be used for research purposes. In addition, **all your information will be strictly confidential**. Please fill out this questionnaire and kindly provide accurate and complete answers. It might take 20 minutes to be completed.

Your participation and time are highly appreciated.

### ❖ Items pertaining to intended effort:

- I would like to take more English courses at my university in the future.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I would like to spend lots of time studying English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- If my teacher were to give the class an optional assignment, I would certainly volunteer to do it.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I am prepared to expend a lot of effort learning English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- Compared to my classmates, I think I study English relatively hard.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

❖ **Items pertaining to the ideal L2 self:**

- I can imagine myself having clear pronunciation in English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I can imagine myself living abroad and having a discussion in English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I can imagine myself speaking English with international friends.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- Whenever I think of my future career, I imagine myself using English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I can imagine myself studying in a university where all my courses are taught in English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

❖ **Items pertaining to the ought-to L2 self:**

- I have to study English because, if I do not, I think my parents will be disappointed with me.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- Studying English is important to me in order to gain the approval of my peers / teachers / family / boss.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I study English because close friends of mine think it is important.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- It will have a negative impact on my life if I don't learn English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- Learning English is necessary because people surrounding me expect me to do so.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

❖ **Items pertaining to the attitudes toward learning English:**

- I like the atmosphere of my English classes.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I always look forward to English classes.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I find English really interesting.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I really enjoy learning English.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

- I would like to have more English lessons at university.

1. Strongly disagree    2. Disagree    3. Slightly disagree    4. Slightly agree    5. Agree    6. Strongly agree

## Appendix C: Post-Intervention Questionnaire (English Version)

Dear Participant,

I would like to thank you for agreeing to participate in the questionnaire. The aim of this questionnaire is to explore your satisfaction with the pronunciation intervention and the intervention outcomes. I would bring to your attention that this study is not meant to assess your English proficiency or serve any other type of examination as your information will only be used for research purposes. In addition, please note that all **your information will be strictly confidential**. Please fill out this questionnaire and kindly provide accurate and complete answers.

- I liked the session of pronunciation instruction.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- I found the length of the pronunciation instruction was appropriate.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- I've observed pronunciation improvements after the instruction.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- I found the pronunciation instruction is more realistic and practical.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- I found the pronunciation instruction suitable for my English level.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- I found materials very helpful in addressing my English pronunciation difficulties.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- I found activities very helpful in improving my English pronunciation.  
1. Strongly disagree    2. Disagree    3. Agree    4. Strongly agree
- What did you find the **most** beneficial about the pronunciation instruction?

---

---

---

- What did you find the **least** beneficial about the pronunciation instruction?

---

---

---

- What amendments would you suggest to make the pronunciation instruction more beneficial?

---

---

---

## Appendix D: Questionnaire for Native English Raters

Dear Rater,

I would like to thank you for agreeing to participate in the assessment task. The aim of this questionnaire is to ask you about the audio-recordings you have assessed. To make terms clear for you, intelligibility means how much you understand from the speaker's speech, while comprehensibility means how easy/difficult it is for you to understand the speaker's speech. Your responses will only be used in my PhD thesis. Hence, **all your information will be strictly confidential**. You need to complete two parts in this questionnaire. First, fill out the demographic background and teaching experience. Second, you will complete a short questionnaire eliciting your perceptions about the audio-recordings you have assessed. This task might take 25 minutes to be completed. Your assistance and time are highly appreciated.

- **Demographic Background**

1. Evaluator ID:

2. Gender: Male                      Female

3. Date of birth (age):

4. Country of birth:

5. What is your native English accent?

---

---

6. How long have you been teaching English and/or Linguistics?

---

---

7. Which Linguistics and teaching English, including pronunciation, courses have you studied or have you taught?

---

---

8. Have you ever taught in an EFL context? If so, please comment on it.

---

---

9. Have you ever taught English to Arabic-speaking students? If so, from which Arabic speaking countries?

---

---

- **Your own perceptions about the intelligibility and comprehensibility assessment task.**

Please complete the following questions that aim to elicit your perceptions about intelligibility and comprehensibility assessments.

**1. Your own perceptions about intelligibility assessments:**

a) Now think of the words and sentences that you perceived as more intelligible. What makes those words and sentences more intelligible than others? Please be as specific as possible.

---

---

b) Now think of the words and sentences that you perceived as less intelligible. What makes those words and sentences less intelligible than others? Please be as specific as possible.

---

---

c) Are there any other issues you would like to comment on about the words and sentences you heard? Any extra information you would like to give us will be greatly appreciated.

---

---

**2. Your own perceptions about comprehensibility assessments:**

a) Now think of the sentences that you perceived as more comprehensible. What makes those sentences more comprehensible than others? Please be as specific as possible.

---

---

b) Now think of the sentences that you perceived as less comprehensible. What makes those sentences less comprehensible than others? Please be as specific as possible.

---

---

c) Are there any other issues you would like to comment on about the sentences you heard and rated? Any extra information you would like to give us will be greatly appreciated.

---

---

**Thank you for completing the task!**

## Appendix E: Segmental Stimuli

- Segmental stimuli are presented in isolated words in carrier sentences to measure speech intelligibility.

- Words in the /p, b/ contrast

/p/	/b/
pack	back
path	bath
pie	bye
lap	lab
cap	cab
rope	robe

*Target sound is on the left, while the contrastive one is on the right.*

- Words in the /v, f/ contrast

/v/	/f/
van	fan
very	ferry
view	few
live	life
leave	leaf
save	safe

*Target sound is on the left, while the contrastive one is on the right.*

- Words in the /tʃ, ʃ/ contrast

/tʃ/	/ʃ/
cheat	sheet
chair	share
cheap	sheep
which	wish
watch	wash
catch	cash

*Target sound is on the left, while the contrastive one is on the right.*

- **Words in the /ŋ, n/ contrast**

/ŋ/	/n/
thing	thin
rang	ran
sung	sun
king	kin
wing	win
rung	run

*Target sound is on the left, while the contrastive one is on the right.*

- **Segmental stimuli are presented in read-aloud sentences to measure speech comprehensibility.**

1. I le**ave** to **visit** my friends.
2. The boy **put** his **cup** on the table.
3. **Ch**ildren **watch** too **much** TV.
4. They **ring** to tell me something**g**.

*The target segmentals are underlined and bold.*

## Appendix F: Suprasegmental Stimuli

- Lexical stress stimuli are presented in isolated words in carrier sentences to measure speech intelligibility (nouns and verbs are differentiated by stress).

Nouns	Verbs
<b><u>pre</u>sent</b>	pre <b><u>sent</u></b>
<b><u>pro</u>ject</b>	proj <b><u>ect</u></b>
<b><u>per</u>mit</b>	per <b><u>mit</u></b>
<b><u>ref</u>und</b>	ref <b><u>und</u></b>
<b><u>pro</u>duce</b>	prod <b><u>uce</u></b>
<b><u>rec</u>ord</b>	rec <b><u>ord</u></b>
<b><u>rec</u>all</b>	rec <b><u>all</u></b>
<b><u>con</u>tract</b>	con <b><u>tract</u></b>
<b><u>con</u>vict</b>	con <b><u>vict</u></b>
<b><u>sus</u>pect</b>	sus <b><u>pect</u></b>

*Stress is placed on the bold and underlined syllable.*

- Lexical stress stimuli are presented in read-aloud sentences to measure speech comprehensibility (nouns and verbs are differentiated by stress).

Pairs	As a verb	As a noun
<b>Record</b>	Click here to <b><u>rec</u>ord</b> yourself.	The university <b><u>rec</u>ord</b> shows you are a good student.
<b>Suspect</b>	I <b><u>sus</u>pect</b> he was here.	You have no <b><u>sus</u>pect</b> about this issue.
<b>Recall</b>	Do you <b><u>rec</u>all</b> your new iPhone?	The <b><u>rec</u>all</b> policy is not clear.
<b>Permit</b>	The laws <b><u>per</u>mit</b> us to stay in this area.	You have to have a new <b><u>per</u>mit</b> .
<b>Produce</b>	Companies that <b><u>pro</u>duce</b> milk.	The farmers have only fresh local <b><u>pro</u>duce</b> .

*Stress is placed on the bold and underlined syllable.*

- Rhythm stimuli are presented in read-aloud sentences to measure speech

**intelligibility and comprehensibility.**

1. (a) I'll <b>ask</b> for volunte <u>ers</u> to <b>help</b> .	(b) I'll <b>ask four</b> volunte <u>ers</u> to <b>help</b> .
2. (a) It's <b>four</b> to <b>seven</b> . (4:00 – 7:00).	(b) It's <b>four two seven</b> . (427).
3. (a) The <b>product</b> is for <b>eyes</b> .	(b) The <b>product</b> is <b>four eyes</b> .
4. (a) Is <b>this</b> to <b>clean</b> ?	(b) Is <b>this too clean</b> ?
5. (a) It's on a <b>street</b> somewhere.	(b) It's on <b>A</b> Street somewhere.

*Bold and underlined words/syllables are stressed.*

- **Final intonation stimuli are presented in read-aloud sentences to measure speech intelligibility and comprehensibility.**

<b>Final falling intonation with certainty (declarative statements)</b> ↘	<b>Final rising intonation with uncertainty (questions)</b> ↗
He's sick.	He's sick?
Ahmed's going to work.	Ahmed's going to work?
You live with your family.	You live with your family?
His father is a teacher.	His father is a teacher?
She's a doctor.	She's a doctor?
We have a meeting.	We have a meeting?
He lost his wallet.	He lost his wallet?
They're here.	They're here?
It's cold.	It's cold?
You're his friend.	You're his friend?

*The arrow indicates a falling or rising pattern of final intonation.*

## Appendix G: Approval Letter to Conduct the Study

الرقم: .....  
التاريخ: / / ١٤١٤ هـ  
المرفقات: .....



المملكة العربية السعودية  
وزارة التعليم  
جامعة الجوف  
عمادة السنة التحضيرية  
رمزه (٥١/٩/٥٠)

### To Whom It May Concern

This is to certify that the Deanship of the Preparatory Year at University of Al-Jouf has no objection in permitting **Mr. Turki Alharbi** to conduct his study in the Preparatory Year.

This letter was given upon **Mr. Alharbi's** request.

Sincerely,

**Dr. Riyadh T. Almatrafi**  
DEAN OF THE PREPARATORY YEAR,  
UNIVERSITY OF AL-JOUF  
28/ 03/ 2018

## **Appendix H: Participant Information Sheet for Students (English Version)**

**Researcher:** I am Turki Alharbi. I am a PhD candidate in the School of Literature, Languages, and Linguistics in the College of Arts and Social Sciences at the Australian National University.

**Project Title:** Improving the speech intelligibility and comprehensibility of English segmental and suprasegmental patterns in the Saudi EFL context.

### **General Outline of the Project:**

- **Description and Methodology:** The current research project aims to examine the effectiveness of explicit pronunciation instruction on improving the speech intelligibility (being understood by native listeners of English) and comprehensibility (being easily understood by native listeners of English) of Saudi EFL learners, especially to determine whether explicit pronunciation instruction plays a positive role in achieving higher levels of intelligibility and comprehensibility in segmentals (individual sounds) or suprasegmentals (prosodic features). In addition, the current research project aims to explore whether non-linguistic factors (age, motivation, and attitude) are relatively correlated with the speech intelligibility and comprehensibility improvements, and to determine whether Saudi EFL learners are able to retain intelligibility and comprehensibility improvements over time.
- **Participants:** 60 Saudi beginner adult EFL learners
- **Use of Data and Feedback:** The results of the current research project will be published in my PhD thesis. As applicable, the results will be published in well-known academic journals and presented in conferences. In addition, the results will be used to design teaching and training materials for English pronunciation in Saudi Arabia. A summary of the results can be sent to you upon request.

### **Participant Involvement:**

**Voluntary Participation & Withdrawal:** Your participation in this project is **voluntary**, and you may, without negative consequences, decline to take part or withdraw from the research without providing an explanation up until data is prepared for publication. You also have a right to remove/edit any part of a recording or any information on demographic information, information sheet, consent form, language background, motivation and attitudes, and post-instruction questionnaires. You may also abstain from answering any question. You can use a variety of computer software applications, such as Audacity, to remove/edit undesired parts in your recordings. Audacity is very secure and easy to use even if you are not familiar with such software applications. In addition, only you and the researcher will remove/edit any information, no third party will be invited. Once you finish recording, you will listen to your recordings and you will have a chance from the date of recording up until data is prepared for publication if you want to edit or remove any information.

- **What does participation in the research entail?** First, you will be asked to read this participant information sheet and sign a consent form. After this stage, you will be asked to have a pretest recording about words and sentences in English. Next, you will have a 3-week pronunciation instruction, 2 hours each week. Once you complete the pronunciation instruction phase, you will have an immediate posttest recording about words and sentences in English. A month after this stage, you will have a delayed posttest recording about words and sentences in English. Then, you will complete short questionnaires for 10 to 15 minutes about your motivation and attitudes towards learning English and your perceptions about the pronunciation instruction phase.
- **Location and Duration:** The study will take place at the Preparatory Year Program at Al-Jouf University and will take about 10 minutes for each battery of pre-, immediate post-, and delayed posttests. In addition, the pronunciation instruction will take three

weeks (2 hours each week). The questionnaires will be completed within 10 to 15 minutes.

- **Risks:** The risks linked with this study are very similar to what you encounter in your English classroom and your everyday life. As you will be asked to record, no personal and sensitive information will be included in the words and sentences you will record. However, you should avoid providing any sensitive information. You will be provided with an opportunity to edit your recording after you finish recording if you desire.
- **Benefits:** This study will help you improve your English pronunciation to confidently speak with more intelligible and comprehensible pronunciation and reach your ultimate goal of learning English. This study will also help policymakers, EFL teachers, and material designers to determine the most effective pronunciation strategies, activities, and techniques for improving speech intelligibility and comprehensibility of EFL learners in Saudi Arabia.

#### **Exclusion Criteria:**

- **Participant Limitation:** To participate in this study, you must be a Saudi beginner-level learner of English in the English Skills 1 (ENG101) class in the Preparatory Year Program (PYP) at Al-Jouf University. You have not been exposed to English in English-speaking countries. You must not have prior pronunciation training/instruction. You must not report speaking and hearing difficulties. You must be a male undergraduate student.

#### **Confidentiality:**

Your confidentiality will be protected as strictly as possible. Regarding your recording, every possible step will be made in order to guarantee your confidentiality. First, your identity will be hidden by using pseudonyms. In addition, you will be given the opportunity to review your

recording and to remove any parts that might risk your confidentiality. Your identity and personal information will also be deidentified in reports and publications stemming from the current research project.

### **Privacy Notice:**

In collecting your personal information within this research, the ANU must comply with the Privacy Act 1988. The ANU Privacy Policy is available at [https://policies.anu.edu.au/ppl/document/ANUP\\_010007](https://policies.anu.edu.au/ppl/document/ANUP_010007) and it contains information about how a person can:

- Access or seek correction to their personal information;
- Complain about a breach of an Australian Privacy Principle by ANU, and how ANU will handle the complaint.

### **Data Storage:**

**Where:** The hard data (such as information sheet, consent form, demographic information and language background) will be stored in locked drawers in my ANU office in order to make sure that data is not accessible to the public. All the electronic data will be stored in my account on iCloud with a highly secure password. Another copy of the electronic data will be stored in a password-protected folder in my own ANU computer. This data will be backed up on my personal Mac computer which is highly secure and requires a password to access. Overall, no one will be able to access the hard and electronic data except the primary researcher.

- **How long:** Once the data collection is complete, your recordings will be stored for five years from the date of my PhD thesis completion. After five years, your data will be archived for future research projects.
- **Destruction of Data:** The original consent form, participant information sheet, demographic information, language background sheets, and the questionnaires will be

securely destroyed at the end of the archive period, which is five years, while the deidentified information, including recordings, will be retained in password-protected cloud storage for further research after those initial five years.

**Queries and Concerns:**

For any queries or concerns, you can contact me at +61 412699582/ +966555015805 and [u6147058@anu.edu.au](mailto:u6147058@anu.edu.au) or my PhD supervisor Dr Manuel Delicado Cantero at [manuel.delicado@anu.edu.au](mailto:manuel.delicado@anu.edu.au). We are pleased to discuss your queries or concerns.

**Ethics Committee Clearance:**

The ethical aspects of this research have been approved by the ANU Human Research Ethics Committee (Protocol 2018/208). If you have any concerns or complaints about how this research has been conducted, please contact:

Ethics Manager

The ANU Human Research Ethics Committee

The Australian National University

Telephone: +61 2 6125 3427

Email: [Human.Ethics.Officer@anu.edu.au](mailto:Human.Ethics.Officer@anu.edu.au)

## Appendix I: Participant Information Sheet for Students (Arabic Version)

### نموذج بيانات المشارك

**الباحث:** تركي نافع الحربي مرشح دكتوراه في قسم اللغويات واللغات والآداب في كلية الآداب والعلوم الاجتماعية في الجامعة الوطنية الأسترالية.

**عنوان الدراسة:** تحسين سلاسة ووضوح نطق الأصوات الإنجليزية المقطعية والفرق مقطعية عند الطلبة السعوديين متعلمي اللغة الإنجليزية كلغة أجنبية.

### الإطار العام للدراسة:

• **الوصف والمنهجية:** تهدف الدراسة الحالية إلى التحقق من مدى فعالية التدريب المباشر على تحسن نطق أصوات اللغة الإنجليزية لدى الطلبة السعوديين متعلمي اللغة الإنجليزية كلغة أجنبية. كما أن الدراسة تهدف إلى قياس العلاقة المحتملة ما بين العوامل الغير لغوية (كالعمر - الدافعية نحو تعلم اللغة- النزعة النفسية والسلوكية اتجاه اللغة الإنجليزية). هذه الدراسة أيضا تهدف إلى قياس مدى استمرارية التحسن المحتمل لنطق أصوات اللغة الإنجليزية على المدى الطويل.

• **المشاركون:** ٦٠ طالب سعودي في المستوى المبتدئ في تعلم اللغة الإنجليزية.

• **استخدام البيانات والتغذية الراجعة:** نتائج هذه الدراسة سوف تنشر في أطروحتي للنيل درجة الدكتوراه. كما من المحتمل أن تنشر هذه النتائج في إحدى المجالات العلمية المحكمة وربما تقدم في المؤتمرات العلمية في الدخال المملكة وخارجها. هذه النتائج من الممكن أن تستخدم في تصميم المناهج التربوية والتدريبية لتعليم نطق اللغة الإنجليزية في المملكة العربية السعودية. نسخة من هذه النتائج قد ترسل إليك وفقاً لطلبكم الكريم.

### طبيعة المشاركة:

**المشاركة التطوعية والانسحاب:** مشاركتك في هذه الدراسة ستكون **تطوعية**. كما يجدر التنويه والتشديد على حقك الكامل في الانسحاب من المشاركة متى أردت دون أن تكون هناك أي عواقب محتملة ودون أن تقدم أي مبررات قبل أن تكون البيانات جاهزة للنشر. كما يحق لك تعديل أو مسح أي جزء من التسجيلات أو أي بيانات في المعلومات الديموغرافية ونموذج التعهد بقبول المشاركة والخلفية اللغوية والاستبيان المخصص لأسئلة الدافعية في تعلم اللغة الإنجليزية والنزعة السلوكية والنفسية اتجاه اللغة الإنجليزية. كما يحق لك الامتناع عن الإجابة على سؤال أو فقرة. سيكون هناك عدة طرق لتعديل أو مسح الأجزاء الغير مرغوبة في التسجيلات، على سبيل المثال برنامج Audacity. يعتبر هذا البرنامج على درجة عالية من الأمان

ويتمتع بسهولة الاستخدام حتى لأولئك الغير معادين على استخدامه. كما أنه لن يكون هناك طرف ثالث يمكنه الاطلاع على البيانات. حالما تنتهي من التسجيلات، سوف تستمع للتسجيلات وسوف يكون لديك الحق من تاريخ التسجيل حتى اعداد البيانات للنشر لتحرير أو مسح أي بيانات تريد.

- **على ماذا تقتضى المشاركة في الدراسة؟** أولاً، ستسأل عن قراءة هذه النموذج وتعبئته ومن ثم توقيع نموذج الموافقة على المشاركة في هذه الدراسة. كمرحلة تسبق فترة التدريب، سيطلب منك تسجيل بعض الكلمات والجمل في اللغة الإنجليزية. بعد ذلك ستتلقى تدريب لمدة ثلاث أسابيع بواقع ساعتين كل أسبوع. حالما تنتهي من التدريب ستقوم في تسجيل مجموعة من الكلمات والجمل في اللغة الإنجليزية. بعد ذلك ستعطي استبيان يهدف لقياس وجهة نظرك حول التدريب الذي تلقينته واستبيان آخر حول دافعتك لتعلم اللغة الإنجليزية ونزعتك السلوكية والنفسية اتجاهها. بعد شهر من تاريخه، سيطلب منك مجدداً تسجيل مجموعة من الكلمات والجمل في اللغة الإنجليزية.
- **المكان والمدة الزمنية:** ستجرى هذه الدراسة في عمادة السنة التحضيرية في جامعة الجوف. ستستغرق ١٠ دقائق لإكمال كل مرحلة تسجيل. كما ستستغرق فترة التدريب ثلاث أسابيع بواقع ساعتين كل أسبوع. ستستغرق الاستبيانات أيضاً ما يقارب ١٥ دقيقة لإتمامها.
- **المخاطر:** من المحتمل أن تواجه بعض المخاطر المشابهة لما تواجهها بحياتك اليومية المعتادة. ولضمان عدم وقوع أي مخاطر، لن يطلب منك تضمين أي معلومات شخصية أو حساسة عندما تسجل مجموعة من الكلمات والجمل في اللغة الإنجليزية. كما ستمنح الفرصة لتعديل أو مسح أي معلومات تريد بعد إتمام كل تسجيل.
- **الفوائد:** تهدف هذا الدراسة إلى تحسين نطقك في اللغة الإنجليزية حت تتمكن من الحديث بكل ثقة وتعزز تقدمك نحو هدفك المنشود في تعلم اللغة الإنجليزية. كما تهدف هذه الدراسة لمساعدة صناع القرار التربوي ومعلمي اللغة الإنجليزية كلغة أجنبية ومصممي مناهج اللغة الإنجليزية في السعودية لتحديد استراتيجيات النطق في اللغة الإنجليزية الفعالة والتدريبات والطرق الناجعة لتحقيق المزيد من الوضوح في نطق الطلبة السعوديين متعلمي اللغة الإنجليزية كلغة أجنبية.

#### المعايير المعتمدة:

- **حدود المشاركة:** للمشاركة بهذه الدراسة، يجب أن تكون طالب في المستوى المبتدئ في اللغة الإنجليزية من مقرر مهارات اللغة الإنجليزية ١ (نجل ١٠١) في عمادة السنة التحضيرية في جامعة الجوف. كما يجب ألا تكون قد درست اللغة الإنجليزية في إحدى الدول الناطقة في اللغة الإنجليزية كلغة أم. كما يجب ألا تكون تلقيت تدريب

على تحسين النطق في اللغة الإنجليزية. كما يجب ألا يكون لديك صعوبات في النطق والسمع. كما يجب أن تكون طالب بكالوريوس.

### الخصوصية:

سيتم كل ما هو ممكن بغية حماية خصوصيتك. بالنظر إلى التسجيلات، ستتخذ كل خطوة ممكنة لضمان خصوصيتك كإخفاء هويتك باستخدام اسم مستعار. علاوة على هذا، سيكون لديك الفرصة لمراجعة تسجيلك بعد الانتهاء لمسح أو تعديل أي جزء قد تجده يشكل تهديد محتمل لخصوصيتك. كما ستخفي هويتك أو أي معلومات شخصية عند نشر هذه الدراسة أو أي غيرها من التقارير البحثية التي قد تُستقى من هذه الدراسة.

### سياسة الخصوصية:

بغية جمع البيانات الشخصية في هذه الدراسة، الجامعة الوطنية الأسترالية ملزمة في الحفاظ على قانون الخصوصية ١٩٨٨. من الممكن معاينة قانون الخصوصية في الجامعة الوطنية الأسترالية على

[https://policies.anu.edu.au/ppl/document/ANUP\\_010007](https://policies.anu.edu.au/ppl/document/ANUP_010007) ، والذي يحتوي على المعلومات التي يستطيع

من خلالها الزائر من:

- الدخول وتصحيح المعلومات الشخصية.
- تقديم شكوى حول أي انتهاك لقانون الخصوصية الأسترالية في الجامعة الوطنية الأسترالية. أيضا من الممكن تتبع سير هذه الشكوى.

### تخزين البيانات:

**المكان:** ستحفظ البيانات الورقية (على سبيل المثال الاستبيانات المعنية في المعلومات الشخصية والديموغرافية) في مكتبي في الجامعة الوطنية الأسترالية لضمان عدم الوصول إليها من قبل الآخرين. أما البيانات الإلكترونية ستحفظ في حسابي على سحابية (أي كلاود) مع رقم سري عالي الأمان. نسخة من البيانات الإلكترونية ستحفظ برقم سري على حاسوبي الخاص في مكتبي في الجامعة الوطنية الأسترالية. ستنسخ البيانات الإلكترونية على حاسوبي المحمول الخاص (ماك بوك إير)، وهو حاسوب عالي الأمان ويتطلب رقم سري للدخول. كل هذا لضمان عدم تمكن أي شخص من الوصول لهذه البيانات.

- **المدة الزمنية:** حالما يكتمل جمع البيانات، تسجيلاتك ستحفظ لخمس سنوات من تاريخ إتمام متطلبات أطروحة الدكتوراه. بعد هذه الفترة، بياناتك ستتم أرشفتها لأغراض المشاريع البحثية المستقبلية.

• **التخلص من البيانات:** النسخة الأصلية من نموذج الإقرار ونموذج بيانات المشارك والبيانات الديموغرافية ونماذج

الخلفية اللغوية والاستبيانات سيتم التخلص منها بشكل سري بعد انتهاء فترة الأرشفة، المقدره بخمس سنوات. بينما

المعلومات التي رُمت، وتشمل التسجيلات، سيحتفظ بها في ملف مزود برقم سري حتى بعد انتهاء فترة الأرشفة.

#### **الاستفسارات والملاحظات:**

إذا كان لديكم أي استفسار أو ملاحظة، من الممكن التواصل معي على الجوال في السعودية (+966555015805) في

أستراليا (+61412699582) أو مراسلتي على بريدي ([u6147058@anu.edu.au](mailto:u6147058@anu.edu.au)) أو بريد مشرفي الدكتور مانويل

ديلكادو كانتيرو ([manuel.delicado@anu.edu.au](mailto:manuel.delicado@anu.edu.au)). من دواعي سرورنا استقبال استفساراتكم وملاحظاتكم.

#### **موافقة لجنة أخلاقيات البحث العلمي:**

حصلت هذه الدراسة على الموافقة من قبل لجنة أخلاقيات البحث العلمي في الجامعة الوطنية الأسترالية بموجب (بروتوكول

٢٠١٨/٢٠٨). إذا كان لديكم أي مرئيات أو شكاوى حول إجراء هذه الدراسة، نرجو التواصل على العنوان التالي:

رئيس أخلاقيات البحث العلمي

لجنة أخلاقيات البحث العلمي في الجامعة الوطنية الأسترالية

الجامعة الوطنية الأسترالية

الهاتف: +61 2 6125 3427

البريد الإلكتروني: [Human.Ethics.Officer@anu.edu.au](mailto:Human.Ethics.Officer@anu.edu.au)

### Appendix J: Written Consent for Students (English Version)

I have read and understood the Participant Information Sheet you have given me about the research project, and I have had any questions and concerns about the project (improving the speech intelligibility and comprehensibility of English segmental and suprasegmental patterns in the Saudi context) addressed to my satisfaction.

I agree to participate in the project YES  NO

I agree to attend pronunciation instruction YES  NO

I agree to my pronunciation being recorded in different tests YES  NO

I agree to audio clips from these recordings being played in future experiments and academic and public settings YES  NO

I agree to be identified in the following way within research outputs:

Full name YES  NO

Pseudonym YES  NO

No attribution YES  NO

I agree to be identified in research outputs via a pseudonym YES  NO

I choose the following pseudonym

---

I would like to receive a copy of research results. My email address is

---

Signature: .....

Date: .....

## Appendix K: Written Consent for Students (Arabic Version)

### نموذج الإقرار للمشاركين

أود الإقرار بأنني أطلعت على نموذج بيانات المشارك حول طبيعة هذه الدراسة، وعليه ليس لدي أي استفسارات أو ملاحظات حول الدراسة المعنونة بـ (تحسين سلاسة ووضوح نطق الأصوات الإنجليزية المقطعية والقوق مقطعية عند الطلبة السعوديين متعلمي اللغة الإنجليزية كلغة أجنبية).

- نعم  لا  أوافق بالمشاركة في هذه الدراسة
- نعم  لا  أوافق بحضور جلسات التدريب
- نعم  لا  أوافق في تسجيل صوتي بمختلف المراحل
- نعم  لا  أوافق في اقتصاص مقاطع صوتية من هذه التسجيلات واستخدامها في التجارب المستقبلية على الصعدين الأكاديمي والعام
- نعم  لا  أوافق على تعريفني في البحوث المستقبلية على هذا النحو:
- نعم  لا  الاسم كامل
- نعم  لا  اسم مستعار
- نعم  لا  دون أي تسمية أو إيعاز
- نعم  لا  أوافق في تعريفني في النتائج البحثية من خلال اسم مستعار
- أختار الاسم المستعار التالي:

---

أود الحصول على نسخة من نتائج هذه الدراسة على بريدي الإلكتروني:

---

التوقيع: .....

التاريخ: .....

## Appendix L: Demographic and Language Background Questionnaire (Arabic Version)

### استبانة المعلومات الديموغرافية واللغوية

عزيزي المشارك، بعد التحية الطيبة:

أود أن أعبر عن بالغ الشكر لقبولك المشاركة بهذه الاستبانة. أن هدف هذه الاستبانة يقوم على جمع معلوماتك الديموغرافية واللغوية. يجدر الإشارة إلى أن هذه الاستبانة لا تعني بأي حال من الأحوال اختبار مستواك اللغوي في اللغة الإنجليزية أو أي نوع من الاختبارات القياسية الأخرى حيث أن المعلومات المراد جمعها ستستخدم للأغراض البحثية فقط. علاوة على هذا، ستكون معلوماتك سرية للغاية. أرجو التفضل في الإجابة على الأسئلة التالية بكل دقة. مع وافر الشكر مرة أخرى.

#### أ. المعلومات الديموغرافية:

1. الاسم أو الكنية (اختياري):
2. الجنس: ذكر أنثى
3. مكان الميلاد:
4. تاريخ الميلاد (العمر):
5. أين نشأت:
6. ما تخصصك في الجامعة:
7. هل تعاني من صعوبات في النطق والسمع؟ إذا الإجابة نعم لا قدر الله، ما المشاكل؟

---

---

#### ب. المعلومات اللغوية:

1. ما لغتك الأم؟ ما لهجتك؟

---

---

2. كم لغة درست أو تعلمت؟ أرجو التفضل بذكرهم؟

---

---

3. في أي عمر بدأت في دراسة اللغة الإنجليزية؟ ما السبب؟

4. هل درست اللغة الإنجليزية قبل الالتحاق في برنامج اللغة الإنجليزية في السنة التحضيرية في جامعة الجوف؟ إذا الإجابة نعم،

أرجو تعبئة الجدول التالي:

مكان الدراسة (على سبيل المثال؛ المدرسة/ معهد خاص/ عن طريق الشبكة العنكبوتية)	مدة الدراسة (على سبيل المثال؛ فصل دراسي)	طبيعة المواد (على سبيل المثال؛ مقدمة في مهارات اللغة الإنجليزية)

5. هل تلقيت من قبل أي تدريب على النطق في اللغة الإنجليزية أو علم الأصوات الإنجليزية؟ إذا الإجابة نعم، أرجو تعبئة الجدول

التالي:

السمة النطقية (على سبيل المثال؛ الأصوات المتحركة/ الأصوات الصامتة/ النبرة/ الإيقاع/ التنغيم/ أخرى)	مدة التدريب (على سبيل المثال؛ أسبوع واحد)	مكان التدريب (على سبيل المثال؛ في الفصل الدراسي/معمل/ عن طريق الشبكة العنكبوتية)

6. هل كان مدرسك ناطق في اللغة الإنجليزية كلغة أم عندما كنت تدرس اللغة الإنجليزية؟ إذا الإجابة نعم، أرجو التفضل في

تعبئة الجدول التالي:

الجنسية	المادة (المهارة)	أي مرحلة دراسية (على سبيل المثال؛ معهد خاص/ المدرسة)
---------	------------------	---


7. هل كنت من قبل في دولة تعتبر فيها اللغة الإنجليزية اللغة الرسمية؟ إذا الإجابة نعم، أرجو التفضل في تعبئة الجدول التالي:

الدولة	مدة الإقامة	سبب الإقامة (دراسة/ سياحة/ زيارة أصدقاء أو أقارب/ علاج)

8. هل تمارس اللغة الإنجليزية خارج القاعة الدراسية؟ إذا الإجابة نعم، أرجو التفضل في تعبئة الجدول التالي:

عدد الساعات في الأسبوع	أمارسها بواسطة (الشبكة العنكبوتية/ التطبيقات/ التلفاز/ أخرى)	السبب (دراسة/ متعة/ مشاهدة الأفلام والأخبار/ أخرى)

9. هل تواجه صعوبات لتكون مفهوم من قبل المدرسين الناطقين في اللغة الإنجليزية كلغة أم في الجامعة؟ إذا الإجابة نعم، أرجو

التفضل في ذكر الصعوبات من وجهة نظرك؟

---



---

\*\*\* وافر الشكر والامتنان على تعبئة الاستبيان \*\*\*

## Appendix M: L2 Motivation Questionnaire (Arabic Version)

### استبانة الدافعية والنزعة نحو تعلم اللغة الإنجليزية

عزيزي المشارك، بعد التحية الطيبة:

أود أن أعبر عن بالغ الشكر لقبولك المشاركة بهذه الاستبانة. أن هدف هذه الاستبانة يقوم على استطلاع مدى العلاقة ما بين الدافعية لديك نحو تعلم اللغة الإنجليزية. كما أود قياس شعورك اتجاه تعلم اللغة الإنجليزية. يجدر الإشارة إلى أن هذه الاستبانة لا تعني بأي حال من الأحوال اختبار مستواك اللغوي في اللغة الإنجليزية أو أي نوع من الاختبارات القياسية الأخرى حيث أن المعلومات المراد جمعها ستستخدم للأغراض البحثية فقط. علاوة على هذا، ستكون معلوماتك سرية للغاية. أرجو التفضل في الإجابة على الأسئلة التالية بكل دقة. مع وافر الشكر مرة أخرى.

• أود أن أخذ المزيد من مواد اللغة الإنجليزية في جامعتي في المستقبل.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أود أن أمضي الكثير من الوقت في دراسة اللغة الإنجليزية.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• إذا ما قرر معلمي إعطاؤنا واجب إضافي، أود بكل تأكيد وبشكل تطوعي القيام به.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أنا على أتم الاستعداد لبذل الكثير من الجهد لتعلم اللغة الإنجليزية.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• مقارنة مع زملائي في الفصل، أعتقد أنني أدرس اللغة الإنجليزية بشكل نسبي بصعوبة.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أتصور بأنني أمتلك نطقاً واضحاً في اللغة الإنجليزية.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أتصور بأنني أعيش في الخارج وأشارك بمحادثة في اللغة الإنجليزية.

- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- أتصور بأنني أتحدث في اللغة الإنجليزية مع طلبة أجنبي.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- متى ما فكرت في مستقبلي الوظيفي، أتصور بأنني أستخدم اللغة الإنجليزية.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- أتصور بأنني أدرس في جامعة حيث تقدم جميع المواد في اللغة الإنجليزية.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- يجب أن أدرس اللغة الإنجليزية بسبب، إذا لم أفعل، أعتقد أنني سأخيب ظني والدي.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- دراسة اللغة الإنجليزية مهمة بالنسبة لي حتى أنال قبول والدي/معلمي/عائلتي/مديري.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- أدرس اللغة الإنجليزية بسبب أن أصدقائي المقربين يعتقدون أنها مهمة.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- سيكون هناك تأثير سلبي على حياتي إذا لم أتعلم اللغة الإنجليزية.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- تعلم اللغة الإنجليزية مهم بسبب أن من حولي يتوقعون مني ذلك.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- أحب أجواء محاضرات اللغة الإنجليزية.
- 1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة
- أتطلع دوماً لمحاضرات اللغة الإنجليزية.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أجد مادة اللغة الإنجليزية ممتعة للغاية.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أستمتع حقا في تعلم اللغة الإنجليزية.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

• أود أخذ المزيد من محاضرات اللغة الإنجليزية في الجامعة.

1 . غير موافق بشدة 2 . غير موافق 3 . غير موافق قليلاً 4 . موافق قليلاً 5 . موافق 6 . موافق بشدة

\*\*\* وافر الشكر والامتنان على تعبئة الاستبيان \*\*\*

## Appendix N: Post-Intervention Questionnaire (Arabic Version)

### استبانة ما بعد التدريب

عزيزي المشارك، بعد التحية الطيبة:

أود أن أعبر عن بالغ الشكر لقبولك المشاركة بهذه الاستبانة. أن هدف هذه الاستبانة يقوم على استطلاع وجهة نظرك حيال مدى الرضا على تدريب النطق الذي تلقيته وأهداف التدريب المنشودة. يجدر الإشارة إلى أن هذه الاستبانة لا تعني بأي حال من الأحوال اختبار مستواك اللغوي في اللغة الإنجليزية أو أي نوع من الاختبارات القياسية الأخرى حيث أن المعلومات المراد جمعها ستستخدم للأغراض البحثية فقط. علاوة على هذا، ستكون معلوماتك سرية للغاية. أرجو التفضل في الإجابة على الأسئلة التالية بكل دقة. مع وافر الشكر مرة أخرى.

• نال استحساني التدريب.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• وجدت مدة التدريب على النطق مناسبة.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• لاحظت تحسن في نطقي بعد التدريب.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• وجدت التدريب على النطق أكثر واقعية وعملي.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• وجدت التدريب على النطق مناسب لمستواي في اللغة الإنجليزية.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• وجدت مواد التدريب نافعة للغاية في معالجة صعوباتي في النطق في اللغة الإنجليزية.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• وجدت التمارين نافعة للغاية لتحسين نطقي في اللغة الإنجليزية.

1 . غير موافق بشدة      2 . غير موافق      3 . موافق      4 . موافق بشدة

• ما الذي وجدته الأكثر فائدة من التدريب؟

---

---

---

---

● ما الذي وجدته الأقل فائدة من التدريب؟

---

---

---

---

● ما التعديلات التي تقترحها لجعل التدريب أكثر فائدة؟

\*\*\* وافر الشكر والامتنان على تعبئة الاستبيان \*\*\*

## Appendix O: Participate in a Paid Research Project

If you are a **native speaker of English**, you are invited to participate in a paid research project. I am Turki Alharbi, a PhD student at the Australian National University (ANU). This research project seeks to ask you to assess L2 speech intelligibility and comprehensibility of English learners. In the intelligibility assessment, you will **listen** to a set of short audio-recordings and you will **identify/transcribe** them. In the comprehensibility assessment, you will **listen** to a different set of audio-recordings and you will be asked to rate the speaker on a 9-point Likert scale. Finally, you will **complete** a short post-assessment survey to elicit your perceptions about what you have heard.

- To participate in this research project, **you must meet the following criteria:**
  1. Native speaker of English.
  2. A teacher of English, with at least five years of experience.
  3. You have not taught Arabic speakers of English.
  4. You have not lived/taught in Arabic-speaking countries.
  5. No hearing difficulties.
- **If you agree to participate**, you will complete the task online at a time and place convenient for you (further instructions will be provided once you agree to participate).
- You will be **paid AUD 50** once you complete all the tasks.
- Feel free to email me if you have further questions or queries: ([u6147058@anu.edu.au](mailto:u6147058@anu.edu.au)).

Your participation and time are highly appreciated.

## Appendix P: Audio-Recordings Assessment Instructions

Thank you for agreeing to participate in my research project!

### Instructions:

- Please read the attached Participant Information Sheet and then sign the Consent Form (let me know if you have any questions).
- There are **9 sessions** as follows:

**First session:** Demographic background questionnaire + intelligibility assessments.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_5tm810g59LiQr89](https://anu.au1.qualtrics.com/jfe/form/SV_5tm810g59LiQr89)

**Second session:** Intelligibility assessments.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_72lfNBJMEqgEKPP](https://anu.au1.qualtrics.com/jfe/form/SV_72lfNBJMEqgEKPP)

**Third session:** Intelligibility assessments.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_1Ik03n9I8d2uhZH](https://anu.au1.qualtrics.com/jfe/form/SV_1Ik03n9I8d2uhZH)

**Fourth session:** Intelligibility assessments + post-assessment questionnaire.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_3gcW5FWXa7u3jzn](https://anu.au1.qualtrics.com/jfe/form/SV_3gcW5FWXa7u3jzn)

**Fifth session:** Comprehensibility assessments + post-assessment questionnaire.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_3RgT1LeI27WlpUV](https://anu.au1.qualtrics.com/jfe/form/SV_3RgT1LeI27WlpUV)

**Six session:** Intelligibility assessments.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_bEqJnwyGodtusNT](https://anu.au1.qualtrics.com/jfe/form/SV_bEqJnwyGodtusNT)

**Seventh session:** Comprehensibility assessments + post-assessment questionnaire.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_eKjIMCoeTKoibM9](https://anu.au1.qualtrics.com/jfe/form/SV_eKjIMCoeTKoibM9)

**Eighth session:** Intelligibility & comprehensibility assessments + post-assessment questionnaire.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_6sMBQREdhuMWY2F](https://anu.au1.qualtrics.com/jfe/form/SV_6sMBQREdhuMWY2F)

**Ninth session:** Intelligibility & comprehensibility assessments + post-assessment questionnaire.

[https://anu.au1.qualtrics.com/jfe/form/SV\\_5oTjsTzuO2xFOwR](https://anu.au1.qualtrics.com/jfe/form/SV_5oTjsTzuO2xFOwR)

- Please listen to each audio file once.
- You can complete the task on your laptop at a time and place convenient for you.
- You can take breaks as much as you like.
- Please avoid closing the page if you want to take a break in order to ensure that your responses will not be lost.
- Please let me know once you complete your participation to arrange the payment.

**Your participation and cooperation are highly appreciated!**

## Appendix Q: Participant Information Sheet for Raters

**Researcher:** I am Turki Alharbi. I am a PhD candidate in the School of Literature, Languages, and Linguistics in the College of Arts and Social Sciences at the Australian National University.

**Project Title:** Improving the speech intelligibility and comprehensibility of English segmental and suprasegmental patterns.

### General Outline of the Project:

- **Description and Methodology:** The current research project aims to examine the effectiveness of explicit pronunciation instruction on improving the speech intelligibility (being understood by native listeners of English) and comprehensibility (being easily understood by native listeners of English) of English learners, especially to determine whether explicit pronunciation instruction plays a positive role in achieving higher levels of intelligibility and comprehensibility in segmentals (individual sounds) or suprasegmentals (prosodic features). In addition, the current research project aims to explore whether non-linguistic factors (age, motivation, and attitude) are relatively correlated with the speech intelligibility and comprehensibility improvements, and to determine whether English learners are able to retain intelligibility and comprehensibility improvements over time.
- **Participants:** 15 native English raters
- **Use of Data and Feedback:** The results of the current research project will be published in my PhD thesis. As applicable, the results will be published in well-known academic journals and presented in conferences. In addition, the results will be used to design teaching and training materials for English pronunciation. A summary of the results can be sent to you upon request.

## **Participant Involvement:**

**Voluntary Participation & Withdrawal:** Your participation in this project is **paid**, and you may, without negative consequences, decline to take part or withdraw from the research without providing an explanation up until data is prepared for publication. You also have a right to remove/edit any part of your demographic information, information sheet, and consent form. You may also abstain from answering any question. In addition, only you and the researcher will remove/edit any information, no third party will be invited. Once you finish assessment, you will have a chance from the date of assessment until data is prepared for publication if you want to edit or remove any information.

- **What does participation in the research entail?** First, you will be asked to read this participant information sheet and sign a consent form. After this stage, you will assess recordings that you will hear. Then, you will complete a short questionnaire to elicit your perceptions about what you have assessed.
- **Location and Duration:** The assessment will take place at a time and place you find comfortable so you can hear the recordings clearly. The assessment will take about 55 minutes. In addition, the questionnaire will be completed within 5 to 10 minutes.
- **Risks:** The risks linked with this assessment are very similar to what you encounter in your everyday life. As you will be asked to assess, no personal and sensitive information will be included in the words and sentences you will assess. However, you should avoid providing any sensitive information. You will be provided with an opportunity to edit your assessment after you finish the assessment if you desire.
- **Benefits:** This study will help learners improve their English pronunciation to confidently speak with more intelligible and comprehensible pronunciation. This study will also help policymakers, English teachers, and material designers to determine the

most effective pronunciation strategies, activities, and techniques for improving speech intelligibility and comprehensibility of English learners.

**Exclusion criteria:**

- **Participant Limitation:** To participate in this study, you must be a native listener of English. You are an English teacher. You have not taught in EFL contexts. You must not report hearing difficulties.

**Confidentiality:**

Your confidentiality will be protected as strictly as possible. Regarding your assessment, every possible step will be made in order to guarantee your confidentiality. First, your identity will be hidden by using pseudonyms. In addition, you will be given the opportunity to review your assessment and to remove any parts that might risk your confidentiality. Your identity and personal information will also be deidentified in reports and publications stemming from the current research project.

**Privacy Notice:**

In collecting your personal information within this research, the ANU must comply with the Privacy Act 1988. The ANU Privacy Policy is available at

[https://policies.anu.edu.au/ppi/document/ANUP\\_010007](https://policies.anu.edu.au/ppi/document/ANUP_010007) and it contains information about

how a person can:

- Access or seek correction to their personal information;
- Complain about a breach of an Australian Privacy Principle by ANU, and how ANU will handle the complaint.

**Data Storage:**

**Where:** The hard data (such as information sheet, consent form, demographic information and language background) will be stored in locked drawers in my ANU office in order to make sure

the data is not accessible to the public. All the electronic data will be stored in my account on iCloud with a highly secure password. Another copy of the electronic data will be stored in a password-protected folder in my own ANU computer. This data will be backed up on my personal Mac computer, which is highly secure and requires a password to access. Overall, no one will be able to access the hard and electronic data except the primary researcher.

- **How long:** Once the data collection is complete, your assessments will be stored for five years from the date of my PhD thesis completion. After five years, your data will be archived for future research projects.
- **Destruction of Data:** The original consent form, participant information sheet, demographic information, language background sheets, and questionnaire will be securely destroyed at the end of the archive period, which is five years, while the deidentified information, including assessments, will be retained in password-protected cloud storage for further research after those initial five years.

### **Queries and Concerns:**

For any queries or concerns, you can contact me at +61 412699582/ +966555015805 and [u6147058@anu.edu.au](mailto:u6147058@anu.edu.au) or my PhD supervisor Dr Manuel Delicado Cantero at [manuel.delicado@anu.edu.au](mailto:manuel.delicado@anu.edu.au). We are pleased to discuss your queries or concerns.

### **Ethics Committee Clearance:**

The ethical aspects of this research have been approved by the ANU Human Research Ethics Committee (Protocol 2018/208). If you have any concerns or complaints about how this research has been conducted, please contact:

Ethics Manager  
The ANU Human Research Ethics Committee  
The Australian National University  
Telephone: +61 2 6125 3427  
Email: [Human.Ethics.Officer@anu.edu.au](mailto:Human.Ethics.Officer@anu.edu.au)

### Appendix R: Written Consent for Raters

I have read and understood the Participant Information Sheet you have given me about the research project, and I have had any questions and concerns about the project (improving the speech intelligibility and comprehensibility of English segmental and suprasegmental patterns in the Saudi context) addressed to my satisfaction.

I agree to participate in the project YES  NO

I agree to assess the given recordings YES  NO

I agree to be identified in the following way within research outputs:

Full name YES  NO

Pseudonym YES  NO

No attribution YES  NO

I agree to be identified in research outputs via a pseudonym YES  NO

I choose the following pseudonym

---

I would like to receive a copy of research results. My email address is

---

Signature: .....

Date: .....