

[DECLARAÇÕES]

Declaro que esta tese é o resultado da minha investigação pessoal e independente. O seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto, nas notas e na bibliografia.

Declaro ainda que tomei conhecimento do Código de Ética da Universidade NOVA de [disponível na intranet](#) e que foram respeitados os seus termos no decorrer do presente trabalho de investigação.



Assinado por: Julie Parker
Mason
Identificação: 8032843653
Data: 2024-09-04 às 13:40:21

O candidato,

Lisboa, de de

Declaro que esta Dissertação se encontra em condições de ser apreciado pelo júri a designar.

Declaro ainda que tomei conhecimento do Código de Ética da Universidade NOVA de Lisboa [disponível na intranet](#).

O(A) orientador(a),

Assinado por: Ana Maria Lavadinho Madeira
Num. de identificação: 07361552
Data: 2024-09-06 09:06:18+01'00'

Lisboa, de de

[DECLARAÇÕES]

Declaro que esta tese é o resultado da minha investigação pessoal e independente. O seu conteúdo é original e todas as fontes consultadas estão devidamente mencionadas no texto, nas notas e na bibliografia.

Declaro ainda que tomei conhecimento do Código de Ética da Universidade NOVA de [disponível na intranet](#) e que foram respeitados os seus termos no decorrer do presente trabalho de investigação.



Assinado por: Julie Parker
Mason
Identificação: 8032843853
Data: 2024-09-04 às 13:40:21

O candidato,

Lisboa, de de

Declaro que esta Dissertação se encontra em condições de ser apreciado pelo júri a designar.

Declaro ainda que tomei conhecimento do Código de Ética da Universidade NOVA de Lisboa [disponível na intranet](#).

O(A) orientador(a),

Assinado por: Susana Mesquita de Deus Correia
Num. de identificação: 11409993
Data: 2024-09-04 17:14:00 +0100

Lisboa, de de

Beyond the Suprasegmental:
Pedagogic Interventions to Aid Native-like Pronunciation in
Advanced Speakers of English as a Foreign Language
(Revised Version)

Julie Parker Mason

Doctoral Thesis in Linguistics
Area: Linguistics and Language Teaching

Feb 2025

Tese apresentada para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Linguística (Especialidade em Linguística e Ensino de Língua), realizada sob a orientação científica de Dra Ana Madeira e Dra Susana Correia

Dedicatória pessoal
To Dr Geoffrey Parker
“What will survive of us is love”
From ‘An Arundel Tomb’ by Philip Larkin

AGRADECIMENTOS

My grateful thanks to everyone who made this work possible. To Ana and Susana for their sage advice throughout the project, to the participants and raters who took part in the study, to Radovan and Zoe for their help with the statistical analysis, to Rob, Steph and Richard for providing native speaker speech samples, to Bruna and Juliano for going through this with me, to my mother for giving me a love of language, to Geoffy, Maddy and Emelye who brought me happiness even when things were not going well, but most of all to Steve, without whose practical help and encouragement none of this would have been possible.

Thank you.

Beyond the Suprasegmental:
Pedagogic Interventions to Aid Native-like Pronunciation in
Advanced Speakers of English as a Foreign Language

Julie Parker Mason

ABSTRACT

The attainment of nativeness in oral production is an area which has received little attention in recent years, due to the prioritising of intelligibility and comprehensibility as the goals of pronunciation instruction. There are strong arguments for focusing on intelligibility in second language (L2) pronunciation instruction, since this must be considered the minimum which learners need in order for their utterances to be understood. However, even if the goal of instruction is intelligibility rather than nativeness, there still needs to be a target accent to use as the model if intelligibility is compromised. As there is evidence that nativeness is strongly implicated in intelligibility and that many learners still wish to work towards nativeness in L2, this research focuses on Standard British English as the target of instruction.

The aim of the current study was to investigate two important issues pertaining to pronunciation modification. Firstly, it aimed to add to the research into phonological self-awareness, which suggests that learners may not be aware of how they sound to others in their L2, but intervention might help them to gain greater self-awareness. Secondly, it aimed to evaluate the relative effectiveness of two different approaches to the teaching of pronunciation: a bottom-up approach versus a top-down approach and to evaluate whether either, or both interventions were superior to the Control Group, who received exposure to the target accent without instruction. The bottom-up approach was structure-led, following a contrastive analysis approach to segmental and suprasegmental instruction, which predicted possible areas of native language (L1) interference for speakers of European Portuguese learning L2 English. The top-down approach involved helping students to gain an awareness of the social and psychological influences on their L2 accent and introduced holistic learning strategies which might help them modify their accent, such as mimicry and role play.

Findings indicated that many participants' phonological self-assessment did not match that of native-speaker raters, but that intervention helped to raise their self-awareness. Findings also indicated that the intervention groups outperformed the Control Group post intervention in that native-speaker mean rating scores showed greater change towards nativeness. However, no significant difference between the two intervention groups was evident.

Keywords: intelligibility, comprehensibility and nativeness; phonological self-awareness; Noticing Hypothesis; Skill Acquisition Theory; bottom up and top-down instruction

RESUMO

O desenvolvimento de competência nativa (conhecida por ‘*nativeness*’) na produção oral é uma área que recebeu pouca atenção nos últimos anos, devido à priorização da inteligibilidade e da compreensibilidade como objetivos da instrução de pronúncia. Há fortes argumentos para ter como foco a inteligibilidade na instrução de pronúncia de segunda língua (L2), uma vez que isso deve ser considerado o mínimo que os alunos precisam de ter para que suas produções sejam compreendidas. No entanto, mesmo que o objetivo da instrução seja a inteligibilidade e não a ‘*nativeness*’, continua a ser necessário haver um sotaque alvo para usar como modelo, se a inteligibilidade for comprometida. Como há evidências de que a ‘*nativeness*’ está fortemente implicada na inteligibilidade e que muitos alunos ainda desejam trabalhar para a natividade em L2, este trabalho concentra-se no inglês britânico padrão como o alvo da instrução.

O objetivo do estudo atual foi investigar duas questões importantes relativas à modificação da pronúncia. Primeiro, teve como objetivo contribuir para a pesquisa sobre autoconsciência fonológica, o que sugere que os alunos podem não estar cientes de como soam para os outros na sua L2, mas a intervenção pode ajudá-los a obter maior consciência. Em segundo lugar, teve como objetivo avaliar a eficácia relativa de duas abordagens diferentes para o ensino da pronúncia: uma abordagem ‘de baixo para cima’ (*bottom-up*) versus uma abordagem ‘de cima para baixo’ (*top-down*) e avaliar se uma ou ambas as intervenções eram superiores ao grupo de controle que foi exposto ao sotaque alvo sem intervenção. A abordagem *bottom-up* foi conduzida pela estrutura, seguindo uma abordagem de análise contrastiva para instrução segmental e suprasegmental, que previu possíveis áreas de interferência da língua materna (L1) para falantes de português europeu a aprender inglês L2. A abordagem *top-down* envolveu ajudar os alunos a ganhar consciência das influências sociais e psicológicas em seu sotaque L2 e introduziu estratégias de aprendizagem holísticas que podem ajudá-los a modificar seu sotaque, como a imitação e a dramatização.

Os resultados indicaram que a autoavaliação fonológica de muitos participantes não correspondeu à dos avaliadores falantes nativos, mas que a intervenção ajudou a aumentar sua autoconsciência. Os resultados também indicaram que os grupos de intervenção superaram o grupo de controle pós-intervenção, pois as pontuações médias de classificação dos falantes nativos mostraram maior mudança em direção à *nativeness*. No entanto, nenhuma diferença significativa entre os dois grupos de intervenção foi evidente.

Palavras chave: inteligibilidade, compreensibilidade e ‘*nativeness*’; autoconsciência fonológica; hipótese de ‘*Noticing*’; teoria de ‘*Skill Acquisition*’; abordagem ‘*bottom-up*’ e ‘*top-down*’

CONTENTS

Introduction	18
Chapter 1: Background to the Study	23
1.1 Definition of Pronunciation and Accent	23
1.2 Native Speaker – Foreign Accent	24
1.3 Standard Accents as the Target of PI	30
1.4 History of Pronunciation Instruction in Europe and Beyond	33
1.5 Goals and Targets in Pronunciation	36
1.5.1 Intelligibility, Comprehensibility and Accentedness	36
1.5.2 Target Accent of PI	37
1.5.3 Student-Centred Goals	40
1.6 Summary – Background to the Study	42
Chapter 2: Factors Influencing Ultimate Attainment in L2 Pronunciation	43
2.1 Age of Commencement of Second Language Learning	43
2.2 First Language Interference	48
2.3 L2 Exposure and Use	52
2.4 Language Aptitude	55
2.4.1. Cognitive Control	56
2.4.1.1 Phonological Short-Term Memory/Phonological Working Memory	57
2.4.1.2 Inhibitory Control	58
2.4.1.3 Attention Control	60
2.4.1.4 Phonological Self-Awareness	61
2.4.1.5 Summary – Cognitive Control	63
2.4.2 Auditory Processing/Phonetic Coding Ability	64
2.4.2 Mimicry Ability	66
2.4.3 Musical Aptitude	68
2.4.4 Summary – Language Aptitude	70
2.5 Gender	71
2.6 Personality	73
2.7 Attitude and Identity	75
2.7.1 Individual Motivation	76
2.7.2 Group Affiliation	77
2.7.3 L2 Learning Experience	78
2.7.4 Summary – Attitude and Identity	79
2.8 Conclusion - Factors Influencing Ultimate Attainment in L2 Pronunciation	80
Chapter 3: Influence of Instruction on L2 Pronunciation Attainment	81
3.1 Theoretical Models of Pronunciation Instruction	81

3.2 Perceptual Training	83
3.2.1 Perception Training Research	83
3.2.2 Phonological Self-Awareness Raising Activities	86
3.3 Structure-Focused Pronunciation Instruction	88
3.3.2 Segmental versus Suprasegmental Instruction	89
3.3.3 Teaching Methodology	91
3.3.3.1 PPP Approach	92
3.3.3.2 Corrective Feedback	93
3.3.3.3 CAPT (Computer Assisted Pronunciation Training)	94
3.3.3.4 Contrastive Analysis	95
3.3.3.5 Oral Reading	96
3.3.4 Summary - Structure-Focused Production Instruction	97
3.4 Top-Down Pronunciation Instruction – Beyond the Suprasegmental	98
3.4.1 Student-centred Learning	99
3.4.2 Mimicry/Shadowing/Echoing/Mirroring/PPR	99
3.4.3 Channeling/Roleplay	101
3.4.4 Voice Quality/Articulatory Settings	102
3.4.5 Summary – Top-down Pronunciation Instruction	103
3.5 Current Attitudes to the Teaching of Pronunciation	104
3.6 Summary - Influence of Instruction on L2 Pronunciation Attainment	105
 Chapter 4: Research Methodology	 107
4.1 Research Questions	107
4.2 Evaluating Pronunciation	109
4.2.1 Who Should Judge: Choosing Raters	109
4.2.1.1 Native Speaker/Non-Native Speaker Raters	109
4.2.1.2 Experienced/Naïve Raters	111
4.2.1.3 Number of Raters	113
4.2.1.4 Age of Raters	114
4.2.1.5 Conclusion – Choosing Raters	114
4.2.2 What to Judge: Choosing Speech Token	114
4.2.2.1 Length of Token	115
4.2.2.2 ‘Range Effects’ in Speech Tokens	116
4.2.2.3 Type of Task to Elicit Speech Token	116
4.2.2.4 Speaker Preparation Time	117
4.2.2.5 Ordering of Tokens	118
4.2.2.6 Appearance of Speakers	118
4.2.2.7 Conclusion – Choosing Speech Token	118
4.2.3 How to Judge: Rating Speech Tokens	119
4.2.3.1 Scalar versus Sliding Scale Ratings	120
4.2.3.2 Rating Questions	121
4.2.3.3 Conclusion – Rating Speech Tokens	125
4.3 Participants	126
4.4 Initial Data Collection and Rating	129

4.5 Instruction Procedure	130
4.5.1 Group 1 - Bottom-up Intervention	130
4.5.2 Group 2 - Top-down Intervention	137
4.5.3 Control Group	145
4.6 Post and Delayed Post-Intervention Collection and Rating of Speech Tokens	145
 Chapter 5: Results	 147
5.1 Question 1: To What Extent Do Learners' Pre-Intervention Assessments of their own Spontaneous Speech Production Correspond to those of NS Raters?	147
5.1.1 Speaker and Rater Judgements Compared	148
5.1.2 Summary of Findings	159
5.1.3 Discussion – Question 1	161
5.1.4 Conclusion – Question 1	162
5.2 Question 2: Will either or both of the interventions to influence pronunciation produce superior results to mere exposure to a native speaker accent in late learners?	163
5.2.1 Inter-rater Reliability	163
5.2.2 Native Speaker Scores	164
5.2.3 Analysis of Overall Measure (Six Measures Combined) Group 1, 2 and Control	165
5.2.4 Analysis of Measures Individually (Group 1, 2 and Control)	168
5.2.4.1 Nativeness	168
5.2.4.2 Segmental/Word-Level Features	171
5.2.4.3 Sentence-Level Features	172
5.2.4.4 Fluency	174
5.2.4.5 Vocabulary	175
5.2.4.6 Accuracy	176
5.2.5 Discussion – Question 2	177
5.2.3 Conclusion – Question 2	179
5.3 Question 3: Will a top-down strategy-led approach bring late learners closer to achieving a native-like accent than a bottom-up structure-led approach?	179
5.3.1 Analysis of Overall Measure (Six Measures Combined) Group 1 and 2	180
5.3.2 Analysis of Measures Individually (Group 1 and 2)	181
5.3.2.1 Nativeness	181
5.3.2.2 Segmental/Word-Level Features	184
5.3.2.3 Sentence-Level Features	186
5.3.2.4 Fluency	189
5.3.2.5 Vocabulary	191
5.3.2.6 Accuracy	193
5.3.4 Discussion – Question 3	195
5.3.5 Conclusion – Question 3	195

5.4 Question 4: Can intervention influence the learners' ability to evaluate their own spontaneous speech production to correspond to that of native speaker raters?	195
5.4.1 End of Course Feedback Questionnaire	196
5.4.1.1 Question 1: Do you feel the course has given you an increased awareness of your accent in English?	196
5.4.1.2 Question 2: Which features of your accent do you have an increased awareness of?	196
5.4.1.3 Question 3: What would be the accent which you would like to have in English?	197
5.4.1.4 Question 4: Do you feel that the course has helped you to change your pronunciation?	198
5.4.1.5 Question 5: If the course helped you to change, what aspects of your pronunciation has it helped you to change?	196
5.4.1.6 Question 6: Which activities were most useful in helping you to work on pronunciation?	201
5.4.1.7 Question 6: Other Comments	203
5.4.2 Discussion – Question 4	204
5.4.3 Conclusion – Question 4	206
5.5 General Discussion of Results	206
5.6 Summary of Findings	215
Conclusion	217
References	223
Appendices	248
Appendix A – CEFR (CV) Descriptors	248
Appendix B – Speech Token Instructions Email	249
Appendix C – Accent Rating Questionnaire	250
Appendix D – Learner Consent Form and Pronunciation Questionnaire	253
Appendix E – End of Course Feedback	258

LIST OF FIGURES

Figure 1	Overall Mean Rater Score: Native Speakers (rounded to 2 decimal places)	164
Figure 2	Mean Rater Score per Measure: Native Speakers (rounded to 2 decimal places)	165
Figure 3	Mean Rater Score per Group: Overall Measure (rounded to 2 decimal places) – Group 1, Group 2 and Control	166
Figure 4	Mean Rating Score per Group: Nativeness (rounded to 2 decimal places) – Group 1, Group 2 and Control	169
Figure 5	Mean Rater Score per Group: Segmental/Word-Level Features (rounded to 2 decimal places) – Group 1, Group 2 and Control	171
Figure 6	Mean Rater Score per Group: Sentence-Level Features (rounded to 2 decimal places) – Group 1, Group 2 and Control	173
Figure 7	Mean Rater Score per Group: Fluency (rounded to 2 decimal places) – Group 1, Group 2 and Control	174
Figure 8	Mean Rater Score per Group: Vocabulary (rounded to 2 decimal places) – Group 1, Group 2 and Control	175
Figure 9	Mean Rater Score per Group: Accuracy (rounded to 2 decimal places) – Group 1, Group 2 and Control	176
Figure 10	Mean Rater Score per Group: Overall Measures (rounded to 2 decimal places) - Figure 3 minus Control	180
Figure 11	Mean Rater Score per Group: Nativeness (rounded to 2 decimal places) - Figure 4 minus Control	182
Figure 12	Mean Rater Score per Group: Segmental/Word-Level Features (rounded to 2 decimal places) - Figure 5 minus Control	184
Figure 13	Mean Rater Scores per Group: Sentence-Level Features (rounded to 2 decimal places) - Figure 6 minus Control	187
Figure 14	Mean Rater Score per Group: Fluency (rounded to 2 decimal places) - Figure 7 minus Control	189
Figure 15	Mean Rater Score per Group: Vocabulary (rounded to 2 decimal places) - Figure 8 minus Control	191
Figure 16	Mean Rater Score per Group: Accuracy (rounded to 2 decimal places) - Figure 9 minus Control	193

LIST OF TABLES

Table 1	Rater Profile	114
Table 2	Participant Profile	128
Table 3	Speaker 1 - Pre-Intervention comparison between speaker and rater accent judgements	148
Table 4	Speaker 2 – Pre-Intervention comparison between speaker and rater accent judgements	149
Table 5	Speaker 3 - Pre-Intervention comparison between speaker and rater accent judgements	140
Table 6	Speaker 4 - Pre-Intervention comparison between speaker and rater accent judgements	150
Table 7	Speaker 5 - Pre-Intervention comparison between speaker and rater accent judgements	151
Table 8	Speaker 6 - Pre-Intervention comparison between speaker and rater accent judgements	152
Table 9	Speaker 7 - Pre-Intervention comparison between speaker and rater accent judgements	153
Table 10	Speaker 8 - Pre-Intervention comparison between speaker and rater accent judgements	153
Table 11	Speaker 9 - Pre-Intervention comparison between speaker and rater accent judgements	154
Table 12	Speaker 10 - Pre-Intervention comparison between speaker and rater accent judgements	154
Table 13	Speaker 11 - Pre-Intervention comparison between speaker and rater accent judgements	155
Table 14	Speaker 12 - Pre-Intervention comparison between speaker and rater accent judgements	155
Table 15	Speaker 13 - Pre-Intervention comparison between speaker and rater accent judgements	156
Table 16	Speaker 14 - Pre-Intervention comparison between speaker and rater accent judgements	157
Table 17	Speaker 15 - Pre-Intervention comparison between speaker and rater accent judgements	158
Table 18	Speaker 16 - Pre-Intervention comparison between speaker and rater accent judgements	158
Table 19	Pre-Intervention Mean Rater Score for Nativeness (1-9) for speakers rating their accent differently from NS raters (rounded to 1 decimal place)	160
Table 20	Mean Rater Score per Speaker: Overall Measure - Group 1, Group 2 and Control (rounded to 2 decimal places)	167
Table 21	Mixed ANOVA: Overall Rater Score - Group 1, Group 2 and Control (rounded to 2 decimal places)	168

Table 22	Mean Rater Score per Speaker: Nativeness - Group 1, Group 2 and Control (rounded to 2 decimal places)	170
Table 23	Mixed ANOVA: Nativeness - Group 1, Group 2 and Control (rounded to 2 decimal places)	171
Table 24	Mixed ANOVA: Segmental/Word-level Features - Group 1, Group 2 and Control (rounded to 2 decimal places)	172
Table 25	Mixed ANOVA: Sentence-level Features - Group 1, Group 2 and Control (rounded to 2 decimal places)	173
Table 26	Mixed ANOVA: Fluency - Group 1, Group 2 and Control - Group 1, Group 2 and Control (rounded to 2 decimal places)	174
Table 27	Mixed ANOVA: Mean Scores for Vocabulary - Group 1, Group 2 and Control (rounded to 2 decimal places)	175
Table 28	Mixed ANOVA: Accuracy - Group 1, Group 2 and Control (rounded to 2 decimal places)	176
Table 29	Age of Participants at Start of Study	178
Table 30	Mixed ANOVA: Overall Measures – Groups 1 and 2 (rounded to 2 decimal places)	181
Table 31	Mean change in Rater Score per Speaker: Nativeness - Groups 1 and 2 (rounded to 2 decimal places)	183
Table 32	Mixed ANOVA: Nativeness – Groups 1 and 2 (rounded to 2 decimal places)	184
Table 33	Mean change in Rater Score per Speaker: Segmental/Word-Level Features – Groups 1 and 2 (rounded to 2 decimal places)	185
Table 34	Mixed ANOVA: Segmental/Word-Level Features – Groups 1 and 2 (rounded to 2 decimal places)	186
Table 35	Mean change in Rater Score per Speaker: Sentence-Level Features – Groups 1 and 2 (rounded to 2 decimal places)	188
Table 36	Mixed ANOVA: Sentence-Level Features – Groups 1 and 2 (rounded to 2 decimal places)	188
Table 37	Mean change in Rater Score per Speaker: Fluency – Groups 1 and 2 (rounded to 2 decimal places)	190
Table 38	Mixed ANOVA: Fluency – Groups 1 and 2 (rounded to 2 decimal places)	190
Table 39	Mean change in Rater Score per Speaker: Vocabulary – Groups 1 and 2 (rounded to 2 decimal places)	192
Table 40	Mixed ANOVA: Vocabulary – Groups 1 and 2 (rounded to 2 decimal places)	192
Table 41	Mean Change in Rater Score per Speaker: Accuracy – Groups 1 and 2 (rounded to 2 decimal places)	194
Table 42	Mixed ANOVA: Accuracy – Groups 1 and 2 (rounded to 2 decimal places)	194
Table 43	Group 1: Which features of your accent do you have an increased awareness of?	197

Table 44	Group 2: Which features of your accent do you have an increased awareness of?	197
Table 45	Group 1: If the course helped you to change, what aspects of your pronunciation has it helped you to change?	198
Table 46	Group 1 Comparison of Responses between Pre-Intervention Questionnaire Question 8 and Post-Intervention End of Course Feedback Question 5	199
Table 47	Group 2: If the course helped you to change, what aspects of your pronunciation has it helped you to change?	200
Table 48	Group 2 Comparison of Responses between Pre-Intervention Questionnaire Question 8 and Post-Intervention End of Course Feedback Question 5	201
Table 49	Group 1 and 2: Which activities were most useful in helping you to work on pronunciation?	202
Table 50	Change in Mean Rater Score from Pre to Delayed Post-Intervention per Measure for Group 1, Group 2 and Group 2 without Speaker 15.	208

LIST OF ABBREVIATIONS

ANOVA	Analysis of Variance
AO	Age of Onset
AoA	Age of Acquisition
CAPT	Computer Assisted Pronunciation Training
CEFR-CV	Common European Framework of Reference for Languages - Companion Volume
CF	Corrective Feedback
CLIL	Content and Language Integrated Learning
CLT	Communicative Language Teaching
CPH	Critical Period Hypothesis
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
EP	European Portuguese
ESL	English as a Second Language
FFI	Form Focused Instruction
FonF	Focus on Form
FonFS	Focus on Forms
FonM	Focus on Meaning
FonP	Focus on Performance
GA	General American
HVPT	High Variability Phonetic Training
ID	Individual Difference
IPA	International Phonetic Alphabet
L1	Native Language
L2	Second Language
L2LP	Second Language Linguistic Perception Model
LoE	length of exposure
MLE	Multicultural London English
NNS	Non-Native Speaker
NS	Native Speaker
PAM-L2	Perceptual Assimilation Model of L2
PI	Pronunciation Instruction
PPR	peer-reviewed pronunciation reading
PTSM	Phonological Short-Term Memory
PWM	Phonological Working Memory
RLS	Reverse Linguistic Stereotyping
RP	Received Pronunciation
SBE	Standard British English
SLA	Second Language Acquisition
SLM	Speech Learning Model
SLM-r	Revised Speech Learning Model
TBL	Task-based Learning
UA	Ultimate Attainment
VOT	Voice Onset Time

INTRODUCTION

While high-level learners of a second language are often able to attain a near-native speaker level of grammatical accuracy and lexical range in oral production, it is very rare for a non-native speaker (NNS) to acquire a native-like accent after a certain age.

The *Critical Period Hypothesis* (CPH) suggested that there was a critical period for language learning in childhood and, if that *window of opportunity* had passed, native-like acquisition was no longer possible and learning would be qualitatively different from, and less successful than, the acquisition of a mother tongue (Lenneberg, 1967). It has been suggested that, for pronunciation, this decline might start at around 5-6 years of age (Long, 1990). However, subsequent research has begun to indicate that the reason why *late learners* - learners who begin to learn a second language after the critical period - generally fail to achieve native-like pronunciation is probably far more complex than this hypothesis suggests (Kinsella & Singleton, 2014; Moyer, 2014; Muñoz & Singleton, 2011). Other factors, such as native language interference (Best, 1994; Escudero, 2009; Flege & Bohn, 2021), quantity and quality of exposure to the L2 (Piske et al., 2001; Uzal et al., 2015), language aptitude (Abrahamsson & Hyltensam, 2008; Sleva & Miyake, 2006), motivation (Moyer, 2014a), degree of phonological self-awareness (Trofimovich et al. 2016; Kennedy & Trofimovich, 2017), identity (Gatbonton et al., 2011; LeVelle & Levis, 2014; Lybeck, 2002; Moyer, 2004) and personality (Hu & Reiterer, 2009; Lewandowski & Jilka, 2019) also appear to be implicated. This has led some to suggest that acquiring a native-like accent, given the right conditions, is possible after the critical period (Birdsong, 2007; Bongaerts et al., 1995; Dollman et al., 2020; Ioup et al. 1994; Moyer, 1999).

While there is an awareness that L2 pronunciation acquisition is influenced by many factors, research seems to be heavily weighted towards the influence of L1 on attainment and the study of de-contextualised phonological forms (Lee et al., 2015; Saito, 2012), which is at odds with the recommendations of important researchers in the field that pronunciation instruction (PI) needs to go beyond structure to look at social and psychological factors, such as personality and identity, which might also have an important role in attainment (Moyer, 2021; Pennington, 2019). Furthermore, as recently as 2018, Darcy bemoaned the fact that, “despite being widely recognized as important, pronunciation does not receive equal attention in the language curriculum” (p.13). This is, at least in part, due to calls for a shift in the goal of PI away from *nativeness* (sounding like a native speaker) to *intelligibility* (getting the message across) and *comprehensibility* (the listener’s perception of difficulty in understanding the message), as proposed by Munro & Derwing (1995) and supported by many in the field (Jenkins, 2000; Levis, 2020; Saito et al., 2017). While this is a laudable change and the beauty and variety of intelligible accents is to be celebrated, it leaves something of a void at the heart of

pronunciation instruction. With learners who are judged to be *unintelligible*, it is unclear which accent to use as the target of instruction, since intelligibility seems to be dependent on the specific profile of speaker and listener (Pérez-Ramón et al., 2022). Furthermore, bearing in mind the current emphasis on student-centred learning (Bernaus, 2010), it is surely the student who should decide what accent they wish to work towards, so teachers need to be prepared to also help those learners who choose nativeness as their goal (Derwing, 2003; Jenkins, 2007; Wach, 2011).

Taking these issues into consideration, this research attempts to take a fresh look at how pronunciation could be addressed in the language classroom. It was inspired by my own experience of teaching English as a Foreign Language (EFL) for over 30 years. Adult classes generally start with a needs analysis: what do the students want to work on in the lessons? What do they feel they needed to improve? Oral fluency, grammar and vocabulary enrichment are always in the mix, sometimes writing, but it is also noticeable how often students ask to work on pronunciation. On one memorable occasion, a student looked me in the eye and said something to the effect of: “will we ever sound like you?” Despite my teaching qualifications and many years of experience, I felt unable to answer this question. Was it even possible for my adult students to change their accents if they wanted to and, if so, what should I be doing to help them?

In an attempt to look at these issues, Chapter 1 starts by looking at what we mean by a native speaker and a foreign accent and discusses which accents are the most useful as the target of pronunciation instruction. Chapter 2 looks at factors affecting L2 pronunciation in late learners, including a discussion of what is meant by the term late learner in the context of pronunciation acquisition (Flege et al., 1999; Granena & Long 2013; Huang, 2014) and a review of the factors affecting L2 pronunciation attainment mentioned above, namely L1 interference, quantity and quality of exposure to the L2, language aptitude, motivation, degree of phonological self-awareness, identity and personality.

Work on phonological self-awareness is of particular interest in this study. In Skill Acquisition Theory, the Dunning-Kruger effect (Kruger & Dunning, 1999; Carter & Dunning, 2008), suggests that learners show a tendency to overestimate their ability and subsequent research has proposed that this may also be the case in pronunciation self-assessment (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al., 2016). If it is the case that learners are unaware what they sound like to listeners, some have proposed that helping learners to gain phonological self-awareness needs to be the first port of call of any PI intervention (Couper, 2015; McGregor & Reed, 2018).

Chapter 3 then looks at theoretical models of learning as they relate to PI, namely, the *Noticing Hypothesis* (Schmidt, 1990, 2001) and *Skill Acquisition Theory* (DeKeyser & Criado, 2013; DeKeyser, 2015). This is followed by an overview of research into PI with a view to establishing which types of

intervention have proven to be more effective in achieving accent modification. This will involve a discussion of research evidence into the role of perception training (Sakai & Moorman, 2018) and interventions to raise learner phonological self-awareness (Couper, 2011; Lappin-Fortin & Rye, 2014; Trofimovich et al., 2016), the relative benefits of segmental versus suprasegmental instruction (Gordon & Darcy, 2016; Saito et al., 2016; Thomson & Derwing, 2015a), the effectiveness of structure-focused versus meaning-focused instruction (Ellis, 2016) and findings regarding structure-focused production instruction (Thomson & Derwing, 2015a; Saito, 2012; Lee et al., 2015). Finally, LaScotte et al. (2021) describe how the research and teaching of L2 PI largely falls into two groups: the ‘bottom-up’ approach, focusing on de-contextualised phonological forms (the ‘structure-focused production instruction’ mentioned above), and the ‘top-down’ approach in which communication and the social context are paramount (Pennington, 2019). So, Chapter 3 closes with a summary of research into top-down interventions (LaScotte et al., 2021; Pennington, 2019).

Chapter 4 gives a detailed description of the current study, which focuses on two important issues pertaining to pronunciation modification. Firstly, it attempts to add to the research into phonological self-awareness suggesting that learners may not be aware of how they sound to others (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al., 2016), and to look into the effectiveness of activities which may help them gain awareness (Couper, 2003; Lappin-Fortin & Rye, 2014; Tsunemoto et al., 2022). Secondly, the study aims to compare a bottom-up, structural approach to PI with a top-down, strategy-led approach in achieving accent modification in the direction of nativeness according to native speaker (NS) rater judgements. It also attempts to ascertain whether either or both are superior to a Control Group, exposed to the target accent but without PI.

The 16 participants in the intervention were all L1 speakers of European Portuguese (EP) who self-reported a level of L2 English of B2 or above, who had first been exposed to English at the age of 5+ and who chose nativeness, in the form of Standard British English (SBE), as their pronunciation goal. It is important to emphasise that the choice of pronunciation goal was not a rejection of intelligibility as a pronunciation goal, but rather a response to the demands of the many students who retain nativeness as the pronunciation goal. Furthermore, since the goal of intelligibility/comprehensibility still requires a target accent, it was hoped that this research would also benefit these learners. Participants were randomly divided into two groups of eight participants (Group 1 and Group 2) who received a 10-week one hour PI intervention. Both groups carried out activities which research has suggested may be effective raising phonological self-awareness: teacher and peer feedback on their accent (Miller & Ng, 1994; Tsunemoto et al., 2022), listening to recordings of their own speech (Couper, 2015) and rater feedback (Lappin-Fortin & Rye, 2014). Group 1 maintained a focus on segmental instruction and worked ‘upwards’ to the study of suprasegmental features of the target. As such, this was a structure-

focused approach, which involved the presentation of features of English identified as problematic for L1 speakers of European Portuguese (Cruz Ferreira, 1995; Rupp, 2018; Sedlackova, 2007; Shepherd, 2001). In contrast, Group 2 received PI which focused on giving students an awareness of the possible factors affecting their current L2 pronunciation and offering strategies to help them overcome factors which might be holding them back from achieving their pronunciation goals. This involved work on the influence on pronunciation of identity (Moreno, 2016; Pennington, 2019) and personality (Pennington & Waxler, 2017), and a focus on holistic pronunciation acquisition strategies such as working on role plays (Floare Bora, 2021; Korkut & Çelik, 2021) and mimicking a target speaker (LaScotte & Tarone, 2019; Moreno, 2016).

Speech samples were collected before and after the 10-week pronunciation intervention and again three months after completion of the course. There was also a Control Group, consisting of B2 level students of EFL at Nova University, Lisbon who received general language instruction, but no specific pronunciation intervention over the 10-week period and provided speech tokens as with the other two groups. Speech tokens were also provided by three English NSs, in order to ensure they were being judged as such by NS raters. All tokens were evaluated by nine NS raters for overall accent, segmental and word level features, suprasegmental features, fluency, vocabulary and accuracy. It was hoped to compare the outcomes of the three groups to inform two specific areas of PI: learner phonological self-awareness and the relative effectiveness of bottom-up versus top-down instruction.

Firstly, to look at pronunciation self-awareness, the rater feedback on the Pre-Intervention participant tokens was compared with learner self-evaluation on the Pronunciation Questionnaire (Appendix D) to analyse whether the participants in the study were evaluating their accents in a similar way to the evaluation of NS raters in the Pre-Intervention Accent Rating Questionnaire (Appendix C). Following the work on self-awareness in pronunciation cited above (Trofimovich et al. 2016; Kennedy & Trofimovich, 2017), it is expected that, at least some of the learners, will not produce descriptions which correspond to those of NS listeners. The same analysis was then repeated after the 10-week PI intervention, to ascertain whether the intervention, which aimed to help learners gain an awareness of their current accent in line with previous research cited above, led to participants and NSs rating in a more similar way. It was expected that receiving pronunciation feedback from the teacher, raters and other learners in the group, along with encouraging speakers to record and listen to their own output, would give participants greater awareness of their current accent in line with previous research findings mentioned above (Couper, *ibid.*; Lappin-Fortin & Rye, *ibid.*; Trofimovich et al., *ibid.*).

Secondly, in order to evaluate the relative merits of a bottom-up versus a top-down approach, the study aimed to look at whether intervention (Groups 1 and 2) produced superior outcomes to exposure (Control Group) in terms of NS rater judgements of nativeness. It was predicted that, in accordance

with other research, intervention would achieve change (Saito, 2012; Lee et al., 2015.; Kennedy & Trofimovich, 2017; Thomson & Derwing, 2015a), but mere exposure would not (Lee et al., *ibid.*; Piske et al. 2001; Saito, 2012; Selinker, 1972).

Finally, it was hoped to compare a structure-led, bottom-up approach (Group 1) with a communicative, top-down approach (Group 2) to ascertain whether one approach proved superior in achieving change towards nativeness in NS rater judgements. It was predicted that the top-down approach, with its more holistic and communicative-led focus, would produce superior results in line with the recommendations of influential researchers in the field (LaScotte et al., 2021; Moyer, 2008; Pennington, 2019). This has been researched in the area of second language listening (Orii-Akita, 2014; Yeldam, 2016) and vocabulary learning (Lee & Yoon, 2019; Moskovsky et al., 2015), but there appears to be no research comparing a bottom-up versus a top-down approach in a general PI programme.

The results of the study are presented in Chapter 5 with a detailed analysis of the findings, a discussion of possible limitations, a summary of the implications of the research for PI and suggestions for future research.

In 2017, Levis wrote the following:

“L2 pronunciation teaching and materials development are full of talented craftspeople who bring divergent, yet under-researched approaches to teaching, and who may show little knowledge of research findings. On the research side, research articles are often written with pedagogical implications that no one applies. The result is that the puzzle pieces that connect research and practice remain too few to see a fuller picture of what works, why it works, and in what situations it works” (p.1)

It is hoped that this research will take a small step towards bridging the gap between research findings and pronunciation teaching.

1. BACKGROUND TO THE STUDY

“Perhaps more than any other aspect, pronunciation is the salient feature of our language competence. It is the lens through which we are viewed in each interaction we have.”

(Celce-Murcia, Brinton & Goodwin, 2010, p. 279)

This chapter starts with a definition of the key concepts of pronunciation, accent, native speaker and foreign accent. This is followed by a discussion of the so-called ‘standard’ accents and an overview of the modern history of pronunciation instruction (PI), including changing attitudes to the place of standard accents as the model. The chapter concludes with a discussion of the motivation behind this research and the methodology used.

1.1 DEFINITION OF PRONUNCIATION AND ACCENT

Darcy (2018) notes that, “spoken language is sound – and sound gives life to grammar and vocabulary. Without sound ... one cannot bring the rest of language to life” (p.13). Of course, spoken language is not just any noise emitted from the mouth. As Rogerson-Revell (2011) points out, “speech sounds are different from other sorts of vocal sounds (vocalizations) because they make regular, meaningful patterns” (p.2). Thus, while a cough or a scream might convey meaning, they would not be considered speech sounds, since they do not form part of the regular patterns of a language. Moyer (2013) notes that *pronunciation* “typically refers to articulation; the place and position of speech organs” (p.10). So, pronunciation can be defined in terms of the production, or *articulation* of speech sounds and how the *articulatory system* of vocal cords, lips, teeth, alveolar ridge, hard palate, soft palate, tongue, jaw and nasal cavity work together with the lungs to produce speech sounds. In the study of the sounds of a particular language, *phonology*, speech sounds are generally divided into *segmental features*, the individual units of speech (phonemes) and features above the segmental level, *prosodic features* including rhythm, pitch, stress, intonation, loudness, duration and related *co-articulatory features* of fluent speech, such as elision, contraction and assimilation. Together, these features are known as *suprasegmental features* of pronunciation.

Pennington and Richards (1986) add a third dimension to their description of pronunciation, that of *voice-setting features*, which they describe as “the tendency of speakers of a particular language to adopt certain habitual positions of articulation in connected speech, resulting in a characteristic voice

quality” (p.209). Voice settings, sometimes referred to as *articulatory settings* or *voice quality*, inform the overall impression listeners have of a given language. Jones & Evans (1995) describe, for example, how French has a “labialized and dentalized quality, while Mandarin Chinese has a nasalized, tense voice quality” (p.245). While voice-setting features could be seen as the broadest level of prosody, they are often treated as a separate from suprasegmentals (see, Pennington & Richards, *ibid.*; Rogerson-Revell, *ibid.*, Jones & Evans, *ibid.*), since the quality of voice is an overarching feature of speech which informs both segmental and suprasegmental production. Pronunciation, then, refers to the production of speech at the segmental, suprasegmental and voice quality level.

While pronunciation focuses on the articulation of speech sounds and sequences, the definition of *accent* focuses on individual and group differences in pronunciation within a given language. Collins and Mees (2009) say the following: “All languages typically exist in a number of different forms. For example, there may be several ways in which the language can be pronounced; these are termed accents” (p.2). Derwing and Munro (2005) expand this definition by describing accent as a “complex aspect of language that affects speakers and listeners in both perception and production, and consequently, in social interaction” (p.385). So, a given accent might be subject to geographical variation depending on the origins of the speaker (e.g. a Liverpool accent, an American accent) and social variation, relating to the cultural and educational background of the speaker, such as *Multicultural London English* (MLE). MLE is a mixture of English and Jamaican, with West Indian and Indian language elements, spoken largely by young, working-class people in Southern England.

Individuals might also vary their accent in accordance with the context in which they are speaking (e.g. a job interview vs. chatting to friends). Furthermore, accent refers, not just to the output of the speaker, but also to its *auditory effect* on the listener.

Thus, while pronunciation refers specifically to speech articulation, accent refers to the unique way in which an individual or a group of people produce the speech sounds of a given language to convey both meaning and social affiliation and the effects of this output on the listener.

1.2 NATIVE SPEAKER – FOREIGN ACCENT

The notion of the *native speaker* (NS) is an important one in second language acquisition (SLA) in general and PI in particular, since, typically, it provides a target for learners, the standard against which language competence is measured. Research suggests that listeners are extremely sensitive to a foreign accent. Flege (1984) found that NS English listeners could accurately identify an NS English accent from French-accented English even after only 30 milliseconds of speech, while Munro et al. (2003)

found that English NS listeners could even detect NNS accents in backwards speech. Anecdotally, anyone who has acquired a language in early childhood is taken to be an NS but the reality proves itself to be more complex. There is disagreement over what is meant by ‘early’ when it comes to language learning and whether there is indeed a golden period for language learning, the critical period, after which native acquisition becomes impossible. This issue is discussed in depth in Section 2.1 below, but there is certainly evidence that early exposure to a language is not the whole story. There are cases of speakers who report a change in what they consider their L1. Kouritzin (1999) describes the case of Lara, a Finnish speaker, who moved to Canada from Finland at the age of 2 but continued to live in a largely Finnish speaking environment until the age of 6 when her family moved to a big city and she was exposed to English. Lara reported that her English began to take over and that, by the age of 18, she was not longer able to speak Finnish and her L1 was now English. Conversely, there are reports of so-called *exceptional learners* who are judged to have native competence despite first learning the L2 in adulthood. For example, Moyer (1999) reports on an English NS who started learning German at the age of 22 and achieved a higher mean rating in the language than one of the NS German controls in her study.

The above research refers to general language competence, but a similar picture emerges when the effect of the CPH on the specific area of accent is studied. There are examples in the literature of young bilingual children who are not judged to have a NS accent, for example Thompson (1991) who reported a perceived foreign accent in the English of 2 Russian NSs who began learning English at the age of 4. Indeed, Piske et al. (2001) in a meta-analysis of studies into the effect of age of acquisition (AoA) on attainment in L2 acquisition conclude, “no study has as yet provided convincing evidence for the claim that L2 speech will automatically be accent-free if it is learned before the age of about 6 years and that it will definitely be foreign-accented if learned after puberty” (p.197).

However, Abrahamsson and Hyltenstam (2008) argue that the apparent existence of exceptional learners may be due to a lack of thoroughness in assessment. For example, in a study of 24 high level American learners of German, Moyer (1999) tested NS ratings of their accents by 4 German NSs in four different conditions: word-list reading, sentence reading, paragraph reading and free speech production. While the first three conditions elicited a considerable number of NS judgements from the raters, only one participant was judged as native in the free speech condition, suggesting that the type of speech token elicited affected the results. In their own study, Abrahamsson and Hyltenstam (ibid.) studied 41 Spanish/Swedish bilinguals, who had been identified as native by NSs of Swedish. The authors carried out what they described as “a battery of 10 highly complex, cognitively demanding tasks and detailed measurements of linguistic performance, representation, and processing” (p.249). The results found that none of the 10 participants over the age of 12 (AoA 13-19 years) and only 3 of

the younger learners (AoA 1-11 years) were judged as performing at the same level as the 15 NSs in the study. In other words, while an L2 speaker might be classed as native or near-native in certain situations or on certain tests, it is very rare that they are classed as such under all conditions. This claim is supported by other research showing that even early bilinguals can be seen to differ from monoglots (Piske et al., *ibid.*; Hyltenstam & Abrahamsson, 2000). Similarly, Sorace (2003) suggests that “even those who are capable of nativelike performance often have knowledge representations that differ considerably from those of native speakers” (p.130), or more recently that, “bilingual speakers may be less efficient at processing structures” (Sorace, 2011, p.17). So, while the speech production of high-achieving NNSs may be comparable to that of NSs, there might be differences in the way this knowledge is stored and/or processed.

In contrast, Davies (2003) in his book *The Native Speaker: Myth and Reality* complains that researchers are going out of their way to devise evermore complex tests in order to try to show that a speaker who learns the L2 after the CPH is different from one who has learned the language in question as an L1 during the critical period and that, therefore, in terms of objective assessments of language attainment, “no test is ever sufficient to demonstrate conclusively that native speaker and non-native speaker are discrete” (p.213).

The problem with separating an NS from an NNS is not with the typical examples: few would argue that a monolingual speaker of English, exposed to English since birth would be defined as an NS, while an L1 Spanish speaker who has attained a B1 level of proficiency in English, with pronunciation heavily influenced by the L1 would be described as an NNS. The problem here lies with what Escudero and Sharwood Smith (2001) describe as “the shadowy borderline that divides native from non-native” (p.277). Part of the difficulty lies with the gamut of accents displayed in those whose L1 is English. In 2011, British singer Cheryl Cole was dropped from the judging panel of an American music competition due to “concerns how American viewers would take to Cole’s Geordie [Northern English] accent” (Plunkett, *Guardian newspaper*: 26 May 2011), which was judged too difficult for American listeners to understand. It would be hard to argue that Cheryl Cole, brought up in Newcastle-upon-Tyne of British parentage, could be termed anything other than a native English speaker, and yet Cole’s production was considered to be so unfamiliar to the target English-speaking audience as to be unintelligible. This suggests that being a NS requires a broad definition to take account of this variety which, at times, leaves NSs being unintelligible to each other. Eubank and Gregg (1999) describe this notion succinctly when they talk of language as “a house with many mansions, and perhaps a few outbuildings as well” (p.66). A definition of ‘native speaker’ requires us to decide which ‘mansions’ and ‘outbuildings’ to include. As Davies (*ibid.*) points out, the tendency is to define NS in terms of NNS:

“To be a native speaker means not being a non-native speaker. Even if I cannot define a native speaker I can define a non-native speaker negatively as someone who is not regarded by him/herself or by native speakers as a native speaker” (p.213).

Davies goes on to note that this distinction means that, under analysis, the concept of an NS is “an emperor without any clothes” (p.213), a social construct rather than a clearly defined reality, identified via the speaker’s self-evaluation and subjective listener judgements rather than through any fundamental properties of the accent per se. Derwing and Munro (2009) offer a similar perspective on foreign accent: “we don’t have a better definition than anyone else, but like many researchers, we have operationalized the construct in terms of listeners’ perceptions of speech” (p.478). The question then moves away from what is an NS to what do listeners *perceive* as an NS?

With this listener-focus in mind, Munro and Derwing (1995b) identify three aspects of the pronunciation of foreign-accented speech which impact the listener: *comprehensibility* – “listeners’ perceptions of difficulty in understanding particular utterances”, *intelligibility* – “the extent to which an utterance is actually understood” and *accentedness* - “how strong the talker’s foreign accent is perceived to be” (p.291), where *foreign accent* is defined as “non-pathological speech that differs in some noticeable respects from native speaker pronunciation norms” (p.289). *Accentedness* can also be defined as being the opposite of *nativeness*, where ‘nativeness’ refers to the degree to which a speaker conforms to NS pronunciation norms. A speaker with a strong foreign accent which strays from pronunciation norms is perceived as an NNS, but what do we mean by ‘strong’ when referring to accent and what are these elusive ‘native speaker pronunciation norms’ which confer nativeness?

Escudero and Sharwood Smith (ibid.) offer a possible answer to this conundrum of finding a definition of ‘native speaker’ by using the work of Rosch (1975, 1978) on Prototype Theory to explain what is going on when listeners make NS judgements. Prototype Theory postulates that a given concept, for example a ‘bird’, would include prototypical examples such as a sparrow or a pigeon, which might be seen as capturing the essence of what we understand as a bird. However, the term ‘bird’ also includes more peripheral examples, such as an ostrich, peripheral because it lacks one or more features which we consider central to the concept of a bird, in this case, the ability to fly. Escudero and Sharwood Smith (ibid.) use this concept to inform the notion of the NS, saying that, as has previously been argued, this is not a black-and-white concept and that there might be certain linguistic features such as accent and grammatical accuracy and non-linguistic features such as nationality which are considered central to the NS prototype by the listener. Equally, other features might be considered more peripheral to the concept, such as linguistic features: range of lexicon, fluency, and non-linguistic features such as appearance and socio-cultural knowledge which could cause uncertainty and might lead one listener to hear *nativeness* where another would not. Thus, as Escudero and Sharwood Smith (ibid.) argue that

“the notion of nativeness should be viewed as a gradient notion associated with a given prototype” (p.283). Using this description of NS, the NS – NNS divide can be seen as a continuum, rather than a finite distinction, with nativeness at one end and accentedness at the other. So, what is the metaphorical ‘sparrow’ of the English accent, what do we mean by a prototype when talking of English pronunciation?

The obvious candidates for the prototype would be the so-called *standard* accents, often referred to as *acrolectal* accents. Finegan (2015) describes a standard variety of a language as one frequently used in “public discourse—newspapers, radio broadcasts, political speeches, college and university lectures, and so on” and which has “undergone a process of standardization, during which it is organized for description in grammars and dictionaries and encoded in such reference works” (p.14). In terms of accent, these are the productions often heard in public discourse and used in the phonetic transcription in dictionaries and teaching materials such as General American (GA), also called North American English (NAE) or Standard American English (SAE), in the United States and Received Pronunciation (RP), often now referred to as Standard Southern British English (SSBE) or Standard British English (SBE), in the UK. Section 1.3 below contains a more detailed discussion of RP and the role of standard accents in general in PI, but in the remainder of this section, the focus will be on their place in defining the prototypical English accents.

Verbeke and Simon (2023) describe a study in which they compare listener judgements of standard, regional and non-native accents for comprehensibility, intelligibility and accentedness, with accentedness being defined as “the degree to which a speaker’s pronunciation deviates from an expected norm or from what a listener expects speaker of English to sound like” (p.3). They asked 33 Dutch speaking English as a Foreign Language (EFL) learners to rate speech samples for comprehensibility, intelligibility and accentedness from 8 speakers with various accents in English: General American English, General British English, Newcastle (regional British accent), Texan (regional American accent), Indian English, Nigerian English and two NNS accents, Chinese and Spanish. For accentedness, participants were asked to use a scalar rating where 1 = no accent and 9 = strong accent) (p.8). Given their definition of accentedness, having ‘no accent’ equated to speaking the ‘expected norm’- what the listener expected an English accent to sound like. The study found that it was the General American English speaker who was evaluated as the least accented, which led the authors to conclude that, “given our definition of accentedness, this suggests that listeners’ expectations of what a speaker of English sound like in terms of accent is primarily a General American English speaker” (p.16). Almost half of participants indicated that the General American English speaker ‘had no accent’ (p.9) but General British English was also given low scores for accentedness compared to the other accents in the study. So, describing an accent as ‘strong’ seems to equate with the distance

from ‘no accent’ - what is expected – the prototype. A ‘weak’ accent is close to the prototype, while a ‘strong’ accent is more distant. The idea of defining the NS in terms of adherence to standard varieties is also the conclusion of Davies (ibid.) who states that, “our searches for the native speaker have returned, again and again, to the question of the standard language ... it makes sense to regard a speaker of one of those codes as a native speaker” (p.214).

The position of a given accent within the prototype would seem to be governed by two important factors, the first of which is the listener’s familiarity with the accent. In a scale from ‘never’ to ‘very often’, 31 of the 33 participants in the Verbeke and Simon (2023) study above self-reported hearing General American English often or very often, 24 made the same judgement for General British English and 7 and 5 for Newcastle English and Texan English respectively. The other accents were self-reported as being heard less frequently. So, to some extent, which accents are considered less accented and therefore more prototypical might be linked to familiarity – the English accents which are heard the most were judged to be closer to the prototype. Texan, Indian, Nigerian, Chinese and Spanish were rated as accented to some degree by every speaker. The interesting outlier was Newcastle English, which received ratings from 1-9 on the rating scale, meaning that the whole scale from ‘no accent’ to ‘strong accent’ was used by participants, so there was little participant agreement on the accentedness of Newcastle English. In terms of Prototype Theory, the Newcastle accent might contain aspects of what Escudero and Sharwood Smith (ibid.) describe as ‘peripheral features’, less central to the notion of an NS English accent and therefore, judged differently by different listeners.

Obviously within familiarity is the notion that, if we are familiar with an accent, we ‘know’ whether it is native or not. While a speaker of British English, will probably accept a Liverpool accent or a Northern Irish accent as being a NS accent, they will not accept a French accent or a Russian accent as ‘native’, since they recognise these accents as coming from parts of the world where English is not an L1. Equally, they might also not accept a Singaporean accent or a Nigerian accent as ‘native’ despite English being the L1 in both countries due to lack of exposure to these accents. In other words, whether a listener accepts an accent as ‘native’ is related to which NS accents they have been exposed to.

The second feature which governs an accent’s position within the prototype is context. Escudero and Sharwood Smith (ibid.) note that, “the clever use of a colloquial expression ... plus a good accent may trigger a perception of nativeness in someone whose performance, to a trained observer, is quite clearly non-native in all other respects” (p.280-281). Similarly, Piller (2002) cites examples of learners who have learned an L2 in adulthood and claim to be mistaken for an NS in certain situations, an ability which Piller labels as ‘*passing*’. Passing is seen by her as a temporary skill, which the learner might use, or might be able to master, in certain situations. Equally there might be situations where an NS is

not judged as such. I well remember an incident which attests to the latter. Born and educated in the UK with English speaking parents, I moved to Portugal at the age of 22. On one occasion, asked for directions by some British tourists visiting Portugal, I duly directed them and was somewhat surprised to be complimented on my excellent English. It would appear that the speaker had ‘profiled’ me as Portuguese and that their judgement was based on purely non-linguistic parameters - because we were in Portugal, maybe because I knew the area well enough to give directions and possibly because of my appearance (dark hair). Had we been in the UK, they would probably not have felt the need to compliment me as the default assumption would have been that I was English, given my accent.

In this section, it has been argued that the terms NS and NNS are not discrete terms and that there is a ‘shadowy borderline’ where the two overlap. These terms can perhaps best be defined in terms of listener judgements and Prototype Theory helps to explain how certain aspects of language performance in general, and pronunciation in particular, might be judged more fundamental to the concept of the NS accent, while other aspects might be seen as more peripheral and therefore less clear-cut in terms of nativeness judgements. It is the standard accents which probably lay claim to being most prototypical, although it has to be borne in mind that, even with standard accents, judgements are influenced by the listener’s familiarity with a given accent, by context and expectations. However, it has been argued that central to the concept of the NS accent is the notion of the standard accents, so the following section will look at changing attitudes towards the standard accent as the *target accent* of PI, in other words, the accent used as the pronunciation model in the classroom.

1.3 STANDARD ACCENTS AS THE TARGET OF PI

In the previous section, the concept of the standard accent was introduced, one which is widely spoken and has been codified in reference materials. Traditionally, in British English, the standard pronunciation has been referred to as *Received Pronunciation* (RP) but the terms *BBC* (British Broadcasting Company) *English* or *Queen’s English* were also used in the past. RP was first codified in the *English Pronouncing Dictionary* (Jones, 1926) and while changes have been introduced to the codification in the ensuing years, there has been much debate about whether this pronunciation standard is representative of the English spoken in modern Britain today: Trudgill and Cheshire (1989) estimate that only 9-12% of the British population speak with this accent and more recently Trudgill and Hannah (2008) downgrade that to a mere 3-5% of the population of England. Others (see Sturiale, 2002; Mompean, 2007) argue for the need to revise the standard in order to more accurately reflect the English spoken in Britain today. Language in general and pronunciation in particular are not static entities, but living modes of communication which are subtly changing all the time and need to be

updated to reflect current usage. Nobody bats an eyelid when, each year, new words are added to the Oxford English Dictionary and obsolete terms drop away and clearly, we need to be doing the same with pronunciation. So, the controversy surrounding RP is partly one of content, a debate around what RP has become in the 21st century, but it is also a question of semantics. The argument is not that we no longer have a standard accent in the UK, but rather what we choose to call this standard accent. The term RP has become linked to an accent no longer widely spoken, while talking of BBC English is also outdated, since the pronunciation of BBC journalists is no longer as one-dimensional as it was in the past, and now reflects the wide and rich variety of accents heard in modern Britain. Equally, it is no longer appropriate to talk about Queen's English, since the British monarch is now a king. In this study, it was decided to follow other researchers (see, for example, Hirst, 1998; Hanani et al., 2013; Lindsey, 2019) by referring to the modern British standard as Standard British English (SBE). However, the term RP will be used when referring to SBE in its historic context.

Sturiale (ibid.) claims that the RP accent in Britain was chosen as the standard pronunciation model for its geographic neutrality, since it is not associated with any particular region of the country. However, socially it is anything but neutral. It is generally seen as a prestige accent, linked to the better-educated members of society, to private schools, to royalty and, as previously mentioned, to the national broadcasting company - the BBC. However, Crystal (2006) points out that it was chosen by the BBC not just because of its social prestige but "because it was the form of pronunciation most likely to be nationally understood" (p.388). Returning again to Cheryl Cole, dropped from an American talent show as her Newcastle accent was judged too difficult for American listeners to understand, a fellow British judge on the show, Simon Cowell, with his SBE accent appeared to have no such problems with making himself understood and continues to work on US television to this day. Research evidence seems to support the claim that standard accents are often judged easier to understand by NSs and NNSs, scoring better on comprehensibility ratings, than non-standard varieties (Major et al., 2002; Tiewtrakul & Fletcher, 2010; Verbeke & Simon, 2023). For example, in the Verbeke and Simon (2023) study mentioned above, the two most frequently heard accents, General American and General British, were considered easier to understand than the non-standard accents: Texan and Newcastle and the less familiar standards: Indian and Nigerian by the L1 Dutch listeners. This has been an important argument in favour of choosing a 'standard' or prototypical accent, such as SBE in the UK or GA in the US as a target accent in PI.

However, the very notion of a 'standard' accent, one which is seen as somehow more "correct", and therefore more prestigious, has been questioned. In the debate around which accent should be the standard for reference materials and language instruction, the work of B B Kachru (1985, 1992) was influential in calling attention to the canonical place of some varieties of English in relation to the

numbers of speakers. He distinguished between three groups of English users: the 'Inner Circle' of native English-speaking countries (e.g. UK, USA, Australia), the 'Outer Circle' of former colonial countries where English has become used as a first language (e.g. Singapore, Nigeria, India) and the 'Expanding Circle' where English is spoken as a second language (e.g. Germany, China, Iran). Kachru questioned the historic dominance of certain standard Inner Circle Englishes as the models for education, and the correspondingly low status of Outer Circle Englishes.

In the ensuing years, Outer Circle Englishes have been increasingly researched and codified as language varieties in their own right, and attention has moved to the Expanding Circle. It has been estimated that only 1 in 4 speakers of English is an NS (see, for example, Crystal, 2003). However, Fergusson (2009) makes the important point that, "the argument from demography (the sheer number of lingua franca users) is not as compelling as it is sometimes presented as being" (p.131) and goes on to explain that this figure does not take into account the frequency or the range of contexts in which these NNSs are using the language. The implication here is that while the monolingual English NSs are using the language in all aspects of their daily lives, this is not the case with L2 English speakers, whose English use is probably restricted to specific domains.

Nevertheless, it is this apparent dominance of NNSs in English language interactions that has led some to criticise what is seen as "a tendency for native speakers to be regarded as custodians over what is acceptable usage" (Seidlhofer, 2005, p.339). This argument has been used by Seidlhofer and others (for example, Jenkins, 2000; Walker, 2010) for the codification and instruction of an English variety which more accurately reflects its use as an international language, in other words, for NNS-NNS interactions. One recent version of such a Global English variety is English as a Lingua Franca (ELF) as proposed by Jenkins (ibid.). However, ELF is not without its problems as a pronunciation model. As a term used to describe NNS-NNS interactions, even Jenkins (2012) seems to recognise the difficulty in encoding it. She has described her early work as "naïve" (p.488) and states that, "ELF communication by its nature entails a substantial element of 'online' variability" (p.490). Furthermore, as Davies and Elder (2006) point out, by encoding ELF, there may be a danger of creating an alternative prestigious standard with the same potential to discriminate against and exclude non-standard speakers as the current standards.

In Section 1.2, it was argued that standard accents are at the very heart of what listeners perceive as an NS accent, and this section has discussed changing attitudes to the place of Inner Circle standard accents and, therefore, their acceptability as a target in PI. The following section looks at how these changing attitudes to standard accents, along with new trends in SLA, have influenced the teaching of pronunciation since the start of the 20th century.

1.4 HISTORY OF PRONUNCIATION INSTRUCTION IN EUROPE AND BEYOND

Prior to the 20th century, Kelly (1969: p. 87) talks about PI as the ‘Cinderella’ of language teaching in that it did not receive the attention granted to its sisters, semantics, lexis and grammar. This changed in the early 20th century with the Reform Movement, which sought to align language teaching with scientific linguistic principles. Reformers like Henry Sweet and Otto Jespersen emphasized the importance of teaching pronunciation through a systematic, phonetic approach. Their approach used the work of the International Phonetic Association, founded in 1886, which introduced the International Phonetic Alphabet (IPA), an alphabet which established a correlation between a symbol and the sound it represented, and which could be used to represent the sounds of any language. The Reform Movement prioritised spoken language and believed that learners should be given phonetic training in the form of the phonetic alphabet, articulatory descriptions and contrastive information to help them replicate the standard sounds of the L2, meaning there was an emphasis on segmental instruction, with the codified standard accents, at that time mainly RP, as the target.

By the mid-20th century, the Reform Movement had spawned more radical versions such as the *Audiolingual Method* based on GA in the US and the *Oral Approach* based on RP in the UK, which were heavily influenced by structural linguistics and Behaviourist Psychology. Pronunciation teaching under this method was characterized by repetitive drills and mimicry, aiming for habit formation through pattern practice. The focus was on accuracy and the elimination of errors, with accuracy defined in terms of adherence to a standard. This reflected a prescriptive, error-correction stance to PI, similar to that of earlier periods but grounded in Behaviourist principles of habit formation. Symptomatic of this approach was the *language laboratory* in which learners would sit alone in a booth with headphones, listening to stretches of L2 speech and repeating what they heard.

The latter part of the 20th century brought a major shift in PI with the popularisation of Communicative Language Teaching (CLT), which emphasized the importance of functional communication over linguistic accuracy. This concept had its origins in sociolinguistics and the works of academics such as Hymes (1972) and Swales (1988) into *communicative competence* – the ability to function in a language, rather than on mastery of language forms. It was brought into the SLA mainstream by approaches such as Task-based Learning (TBL) (Prabhu, 1987; Willis, 1996), which advocated the teaching of language via authentic communication fostered by group tasks made up of goal-oriented activities such as problem-solving or decision-making tasks and Content and Language Integrated Learning (CLIL) (Coyle, 2007; Frigols & Marsh, 2007) which advocated the teaching of class subjects such as history or mathematics through the medium of the L2. In TBL and CLIL, the primary focus is

on using the L2 to achieve practical goals, rather than on language accuracy. In the field of pronunciation, a similar focus on CLT has been reflected in the work of writers such as Celce-Murcia (1983, 1987) and Pica (1995), who encouraged the teaching of pronunciation through communicative activities such as contextualized practice, role-playing, and real-life communication tasks and the expansion of teaching to include a focus of suprasegmental features of production, such as intonation and stress.

Synonymous with the move from structural accuracy to communicative competence was a move from nativeness to intelligibility/comprehensibility as the goal of PI - if the speaker of the L2 is achieving the goals of intelligibility and comprehensibility and getting their message across, notions of ‘accuracy’ and adherence to a standard target are no longer relevant. Important researchers in the field of PI have called for this change, such as Saito et al. (2017), who note that, “it is important to set realistic instructional goals for learners, prioritizing comprehensibility ... over linguistic nativeness” (p.2). and Levis (2020) who talks of the issue in terms of the Intelligibility Principle and the Nativeness Principle and argues that “it is time to embrace the Intelligibility Principle and consign the Nativeness Principle to the past” (p.14).

This move from a structuralist approach to a communicative approach was also mirrored in the discussion around a Focus on Forms approach (FonFS) versus a Focus on Form approach (FonF). FonFS describes intervention where the teacher explicitly focuses on a discrete area of language, often *prescriptively* in that they predict what the learner needs, which is presented and practised in the classroom, described by Ellis (2016) as “a structure-based approach ... where linguistic forms are taught directly and explicitly” (p.405). In contrast, FonF describes communicative, meaning-focused instruction, where a focus on language forms is incidental and in response to learner needs, often called *diagnostic* or *descriptive* instruction. Doughty and Williams (1998) state that, “the fundamental assumption of focus-on-form instruction is that meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across” (p.4).

Other researchers (Jones, 1997; Morley, 1996; Pennington, 1989) conceptualised this as a move away from the *bottom-up* approach, structure-led and focusing on de-contextualised phonological forms to a ‘top-down’ approach in which communication is paramount. Pennington (2019a) describes how, in this approach, a *top-down* pronunciation curriculum would start at the level of broad social context, suprasegmentals, and related behaviours, then systematically progress ‘downwards’ to the segmental level” (p.381). Thus, top-down PI aims to give the learner an awareness of psycho-social aspects of speech production such as identity (Moreno, 2016; Pennington, 2019) and personality (Pennington & Waxler, 2017) and to focus on global pronunciation acquisition strategies such role plays (Floare Bora,

2021; Korkut & Çelik, 2021) and mimicking a target speaker (LaScotte & Tarone, 2019; Moreno, *ibid.*).

A possible outlier to the trend towards a communicative approach is the field of speech perception which, based on theories of language acquisition such as Flege's (1995) *Speech Learning Model (SLM)*, and Best's (1995) *Perceptual Assimilation Model of L2 (PAM-L2)*, proposes that perception precedes production and that, if an L2 phoneme is perceived as similar to a learner's existing L1 sound, the learner might fail to recognise it as a new speech sound and instead integrate the new L2 sound within an existing L1 category. Researchers have had considerable success in training paradigms which help the learner to 'hear' the difference between sounds present in the L2, but which the researchers predict will not be perceived as distinct due to L1 interference (Rato et al., 2014) and in some cases, help learners to then produce sounds which more accurately represent NS norms (Bradlow et al., 1997; Sakai & Moorman, 2018). This line of research has helped to inform our knowledge of how L1 and L2 sounds are categorised and perceived, but in terms of PI, it represents a return to a more structuralist approach, assuming nativeness as the goal and using standard accents as the target, which would appear to run counter to current attitudes.

A further area of pronunciation research which runs at a tangent to the discussion of structural accuracy versus communicative competence, but important nonetheless in any discussion of PI, is the issue of *pronunciation self-awareness*. Coming from the field of skill acquisition, much attention has been paid to the *Dunning Kruger effect* – the oft observed discrepancy between how people judge their own abilities compared with how others judge them (Kruger & Dunning, 1999; Carter & Dunning, 2008). More recently, this discrepancy has also been observed in L2 learner judgements of their own pronunciation (Kennedy & Trofimovich, 2017; Lefkowitz & Hedgcock, 2002; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al. 2016). Trofimovich et al. (*ibid.*) tested 134 L2 speakers of English from different language backgrounds on a picture narrative task, and their own judgements of their nativeness and comprehensibility were compared to the ratings of 3 English NSs. The authors noted that, "only about a third of all L2 speakers were fully calibrated with listeners in their self-assessment" and that, "L2 speakers' rating of their accent and comprehensibility correspond only weakly to listeners' judgements" (p.133). It could be, therefore, that learners are failing to achieve pronunciation goals because they are unaware of their current level of success and mistakenly believe that they have already attained their desired level.

Broadly speaking, however, this section has described how, over the last century or so, PI has moved away from a structuralist, FonFS approach with nativeness as the target, to a communicative, FonF approach with intelligibility/comprehensibility as the priority.

However, this change has also created a problem at the very heart of PI. Derwing and Munro (2015) point out that the move away from formal accuracy led to “a decrease in emphasis on pronunciation in CLT classrooms” (p.176). This decline in PI in the language classroom has been noted by other researchers (Darcy, 2018; Foote et al., 2011; Levis & Wu, 2018), so the next section looks at possible reasons for this decrease in emphasis and offers a possible way forward for PI.

1.5 GOALS AND TARGETS IN PRONUNCIATION INSTRUCTION

In the previous section, it was argued that the dominance of communicative competence in SLA instruction has led to a move away from nativeness as the goal of PI in favour of intelligibility/comprehensibility. However, there are some obstacles to this choice and these will be addressed in the following sections.

1.5.1 Intelligibility, Comprehensibility and Accentedness

Firstly, research into the relationship between these three aspects of accent perception, intelligibility, comprehensibility and accentedness, has tended to agree that they are not mutually exclusive. In their study of NS judgements of Mandarin-accented English, Munro and Derwing (1995b) suggested that “most listeners showed a correlation between intelligibility and perceived comprehensibility and between perceived comprehensibility and accent” (p.302) but that “far fewer listeners showed a significant correlation between intelligibility and accent” (p.303). This suggests that it is possible to be intelligible, while still having heavily accented speech, but that degree of accentedness appears to negatively correlate with ease of understanding. Julkowska and Cebrian (2015) came to a similar conclusion in a study of Polish-accented speech in English as rated by English NSs, Spanish NSs and Polish NSs. They noted that, “the strength of the correlation between accentedness and comprehensibility ratings was moderate to high” (p.22) but “weakest correlations involved the comparison between accentedness ratings and intelligibility scores” (p.22). Other studies have found that accented L2 speech may be more intelligible to listeners who share the L1 of the L2 speaker (Bent & Bradlow, 2003; Kennedy & Trofimovich, 2008). Thus, there appears to be only a weak relationship, but a relationship nonetheless, between accentedness and intelligibility and intelligibility in the L2 may be superior when the listener and speaker share an L1.

Furthermore, both the Derwing and Munro (*ibid.*) study and the Julkowska and Cebrian (*ibid.*) study in the previous paragraph attest to a stronger overlap between comprehensibility and accentedness, which has also been noted by other researchers (see Saito et al., 2017; Thomson, 2018a; Trofimovic & Isaacs, 2012)). However, a study by Appel et al. (2019) into L1 English higher education students' ratings of L2 English higher education students' speech goes further by concluding that comprehensibility and nativeness "were seen as largely indistinguishable in raters' assessments" (p.29). It is important to note that the researchers specifically focussed on lexical variation in the samples, but their finding nevertheless suggests that, in some respects at least, comprehensibility and accentedness are judged as synonymous by listeners. While most of the research into this link has focussed on L2 English, it has also been proposed that the connection between comprehensibility and accentedness is closer in tonal languages. Freeborn and Rogers (2019), in research which used Mandarin NSs to rate L2 Mandarin speakers, concluded that, "concepts of accentedness and comprehensibility may be more intricately linked in lexical tone languages such as Mandarin, in comparison to nontonal languages" (p.75).

Thus, it would appear that accentedness and comprehensibility are relatively strongly related, while accentedness and intelligibility are more weakly related, but related nonetheless. This means that it is wrong to think of PI as a choice between intelligibility and comprehensibility on the one hand and nativeness on the other, since these concepts would appear to be linked. In other words, part of what constitutes comprehensibility, and to a lesser extent intelligibility, is nativeness. That is not to say that a strong foreign accent will inevitably cause communication problems but, it relates back to the discussion in Section 1.3 above, which argued that standard accents are fundamental to the prototype of what listeners 'expect' an accent to sound like and, in this section, it has been argued that straying too far from the prototype may cause comprehensibility and, to a lesser extent, intelligibility issues.

In this section it has been argued that intelligibility/comprehensibility and nativeness are not polar opposites, they are somehow related, so being intelligible and comprehensible would appear to involve some degree of nativeness. More research needs to be done to identify which aspects of nativeness are fostering intelligibility/comprehensibility, but there is a second problem with choosing the latter as the goal of PI and this will be the subject of the following section.

1.5.2 Target Accent of PI

A central issue for those who prioritise intelligibility/comprehensibility as the goal of PI, is the target of instruction. Gilbert (2010) says the following:

"There is a paradox connected with the notion of "communicative"-ness. The communicative approach tended to downplay the importance of accuracy in general, but somehow overlooked the

fact that pronunciation is an immediate barrier to communication unless it has a certain degree of accuracy, in the sense of conformity to some recognised/recognisable system (2009, personal correspondence with Jonathan Marks)” (p.2).

The point is, even if intelligibility is the goal of PI, there still has to be a target to work towards. If a speaker is deemed unintelligible, there must be some way of ‘correcting’ that and of offering an intelligible alternative. Lewis and Deterding (2018) sum up the centrality of this problem when they state, “if there is no clear model to refer to, how do we give guidance to students about how to improve their pronunciation, and how do we determine what needs to be fixed in order to enhance intelligibility?” (p.161). Even those who champion intelligibility recognise this need for a target, such as Pennington (1996), who points out that, “where the people with whom the learner will need to interact are native speakers of English or native speakers of languages different from that of the learner’s native language, it is necessary to aim towards generalized norms for phonology” (p.240).

This issue is not just be an intellectual storm in a teacup but may be a matter of life and death. If English, as is often the case, is being used as the international language of communication in many domains of life, it is important that speakers have some degree of shared understanding of English in general and pronunciation in particular in order to be mutually intelligible. Tiewtrakul and Fletcher (2009) in a review of air-traffic control – pilot communications conclude that, “communication errors, defined by incidents of pilots not understanding, occur significantly more often when speakers are both non-native English” (p.229), while research by Lindhout et al. (2009) estimated that language problems account for 5% to 10% of all accidents in the chemical industry.

Complications surrounding the choice of target accent in PI are reflected in the influential Common European Framework of Reference for Languages (CEFR), a tool devised by the Council of Europe to provide standardisation of the assessment of language proficiency within Europe. Kemmel et al. (2023) note how the first version the CEFR (Council of Europe, 2001) “has been understood by a wide audience to uphold the notion of the native speaker” (p.67). This is certainly evident in terms of pronunciation with, for example, the phonological control descriptor for A1 being “pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by native speakers used to dealing with speakers of his/her language group” and for B1 being, “pronunciation is clearly intelligible even if a foreign accent is sometimes evident” (p.117). This stance was clearly at odds with trends in PI which, as previously mentioned, was moving away from championing the NS as the arbiter of correctness and championing intelligibility as the goal.

However, in 2020 a revised version was produced, the CEFR Companion Volume (CEFR-CV) in which it is recognised that, “the phonology scale was the least successful of those calibrated in the original

research” (Council of Europe, 2020, p.134), and that, “it appeared more appropriate to focus on intelligibility as the primary construct in phonological control” (p.47). The B1 phonological descriptor scales are now divided into 3 sections: overall phonological control, sound recognition and articulation, and prosodic features. References to the native speaker and foreign accent are now gone so that, for example the B1 overall phonological control descriptor now reads, “pronunciation is generally intelligible: can approximate intonation and stress at both utterance and word levels” (p.136) and the B1 descriptor for the new category of Sound Recognition and Articulation adds, “is generally intelligible throughout, despite regular mispronunciation of individual sounds and words he/she is less familiar with” (p.136). While a move away from NS norms as the target is admirable in its intentions, it leaves something of a void at the heart of pronunciation teaching. If learners are required to ‘approximate intonation’, which intonation should they approximate? And how are we to interpret ‘mispronunciation of individual sounds’ - how can you ‘mispronounce’ if there is no ‘correct’ pronunciation with which to compare? The Phonological Scale Revision Process Report sheds some light on the thinking behind these descriptors by stating, “the Assessment [of pronunciation] needs to be modulated according to different factors like the goals of the instructor, the learner, the curriculum and the context” (p.16). There might, for example, be a difference between EFL, where English is taught as a foreign language in a non-English speaking country and ESL (English as a Second Language), where English is being taught as an L2 in an English-speaking country. Thus, an L1 Portuguese speaker in an EFL class in Portugal might have different pronunciation goals to an ESL learner in Canada, just as the teacher and the curriculum might have different objectives in the two contexts. This focus on learner goals and context reflects recent trends in the field of pronunciation research, where the emphasis has moved away from the imposition of one accent, prestige or otherwise, upon the learners, to the offering of a choice (Jenkins, 2007; Derwing & Munro, 2009; Moyer, 2015). The learner, it is argued, needs to be made aware of the issues surrounding accent modification, in order to be able to make their own informed choices about the goals of their own learning, while teachers “should help them to set realistic goals on the basis of current research findings” (Derwing & Munro, 2005, p.384).

These researchers are right to argue that the choice of goal in PI needs to be negotiated depending on context, but whatever the goal is chosen, there still has to be a target, a model to work towards when goals are not being met. As previously discussed, the obvious choice of model would be the standard accents since these appear to be the most prototypical accents and most widely understood. Even the strongest advocates of intelligibility as the goal of pronunciation teaching, recognise this need for a target accent. For example, Celce-Murica et al.(2010), in their highly influential *Pronunciation Teaching*, champion intelligibility when they say, “a more modest and realistic goal is to enable learners to surpass the threshold level so that their pronunciation will not detract from the ability to

communicate” (p.9) and yet their seminal text is unashamedly focussed on standard North American English because, “it is the target variety of the many ESL students living, studying, or working in North America” (p.11). The same can be said for other respected PI materials such as *Clear Speech* (Gilbert, 2012) and *Beyond Repeat After Me* (Yoshida, 2016) which use North American English as the target and *English Phonology and Pronunciation Teaching* (Rogerson-Revell, *ibid.*) which focuses on British English.

In this section it has been argued that, even if the *goal* of pronunciation instruction is intelligibility/comprehensibility, there still has to be a *target* accent, which provides the model for correction if intelligibility/comprehensibility is compromised. While the CEFR-CV leaves the choice of target accent open, respected authors in the field have reverted to standard accents as the target. It has also been shown that the CEFR-CV and prominent researchers are calling for learners to be involved in making decisions about their own learning, so the following section looks at what research has shown about learner goals and targets.

1.5.3 Student-Centred Goals

In Section 1.4, it was stated that many researchers are now calling for intelligibility to replace nativeness as the goal of PI, but Thomson (2018a) comments that, “while we have rightly idealized intelligibility and to a lesser extent comprehensibility as the goals of instruction, we must also accept that we cannot control the reactions of all listeners, and that for some L2 English learners, in some contexts, a demand for acceptability may trump our idealized standards” (p.15). The term *acceptability* is a reference to the work of Szpyra-Kozłowska (2014), who describes *acceptability* as the “amount of irritation caused by a given accent” (p.76). While this may have something in common with comprehensibility, there is clearly a difference between describing an utterance as difficult to understand and describing it as irritating. Thomson is pointing out that there may be reasons, such as issues of acceptability, why some students prefer to choose nativeness over intelligibility.

Derwing & Munro (2015, p.2) add another dimension to this phenomenon with the term *credibility*, which they define as the degree to which accents are positively or negatively received by the listener. Accented speech may, of course, have a positive side. Derwing and Munro (2009) point out that, “it can signal to interlocutors that they may need modified input” (p.484) and there is also mention in the literature of those who choose to speak with a foreign accent. A famous example is the conductor Leopold Stokowski who, despite being born and bred in London, was reported to speak with an Eastern European accent in English, the suggestion being that this gave him an air of sophistication (reported in Derwing & Munro, *ibid.*). However, to claim that a foreign accent is universally well received by the

listener would be untrue. Thomson (ibid.) puts this bluntly when he writes, “if speech is intelligible, it should be acceptable. In fact, this is not the case” (p.14). There is a growing body of research into this *accent prejudice*, suggesting that an utterance spoken in an NNS accent may be negatively socially evaluated, may be perceived as less trustworthy and may even subject the speaker to negative discrimination. This area often involves NNSs self-reported reflections on reactions to their accent (see, for example, Munro et al., 2006; Harrison, 2012), but negative discrimination has also been recorded in listener judgements of NNS speech. Lev-Ari and Keysar (2010) found that trivia statements such as “a giraffe can go without water longer than a camel can” were rated by NSs as “less truthful when spoken by non-native speakers” (p.1095), even when raters knew that the statement had been provided by the experimenter. The authors conclude that this provides evidence that NNSs are being negatively evaluated because of their accent and that, “accent might reduce the credibility of non-native job seeker, eyewitnesses, reporters or news anchors” (p.1095).

Issues of acceptability, credibility and accent prejudice might go some way to explaining why interviews with learners have repeatedly shown that many learners continue to strive for nativeness as their ultimate pronunciation goal. For example, Derwing (2003) in a study of 100 adult ESL learners in Canada found that 95% said they wanted to sound like an NS. Similar findings are reported by Jenkins (2007) who, following interviews with 17 NNS teachers of English from China, Italy, Korea, Malaysia, Poland, Spain, Taiwan and Ukraine states that, “one of the most prevalent beliefs [of the teachers], and one that was articulated by every one of the participants in one way or another was that an NS English accent is ‘good’ and an NNS ‘bad’” (p.209). Wach (2011) reports on a questionnaire study involving 234 Polish university students asked for their opinions on NSs and ELF norms as the target of pronunciation instruction. She concludes, “the findings obtained in the present study clearly indicate that the studied population of English majors, as a whole, expressed a strong preference for acquiring NS pronunciation as opposed to ELF pronunciation levels (p.259).

It is also possible that students and teachers have different attitudes to the issue. Timmis (2002) surveyed 400 students from 14 countries and 180 teachers from 45 countries using questionnaires to assess attitudes to the use of NS norms in EFL/ESL. In the pronunciation section of the questionnaire, students were asked whether they would “prefer to be like Student A or Student B”, where Student A says, “I can pronounce English just like a native speaker now. Sometimes people think I am a native speaker” and Student B says, “I can pronounce English clearly now. Native speakers and non-native speakers understand me wherever I go, but I still have the accent of my country” (p.242). Timmis found that 67% of students chose Student A and 32% opted for Student B. It is unclear how the other 1% reacted. However, when he posed the same question to teachers, giving them the added option of ‘no preference’, 27% chose Student A, 39% Student B and 34% showed no preference. Timmis concluded,

“It appears then that there is a greater tendency among teachers than among students to regard ‘accented intelligibility’ as the most desirable outcome” (p.243) and on whether the native-speaker model should be the target for learners, he says, “there is no reason why it should, but it may be that some students want it to be, even when we least expect it (p.249)”.

There could be several reasons why many learners choose nativeness over intelligibility as their goal. It might be *instrumental*, a desire for the perceived or real positive benefits of speaking the L2 with what listeners adjudge to be NS levels of proficiency. This is often discussed in terms of better job prospects, but it might also be linked to the concerns about credibility and acceptability as discussed above. However, the motivation might also be *integrative*, so that, fuelled by a positive attitude towards the L2 language and culture, learners strive to become part of the culture they so admire.

These issues might go some way to explaining why the Nativeness Principle remains the gold standard for many students. While those who champion the Intelligibility Principle clearly have the interests of the student in mind, it would appear to go against what many learners are aiming for. As Derwing and Munro (2009) point out, “in these days of learner-centred curricula, it seems ironic that some authorities advocate the opposite of what the students want” (p.485).

1.6 Summary – Background to the Study

In this section, it has been argued that the terms NS and NNS are not discrete, but are overlapping concepts. Prototype Theory helps to explain how certain accents might be considered closer to the prototype, what listeners expect a given accent to sound like and standard accents are usually considered to be the most prototypical. Attitudes to the target accent of pronunciation instruction have changed over the decades with a move from away from nativeness, the adherence to an NS model, to intelligibility and comprehensibility as the goal. However, it has been argued that nativeness, intelligibility and comprehensibility are not discrete and that there is a certain degree of overlap between the terms. Secondly, it has been argued that there needs to be a target accent in PI, even if intelligibility/comprehensibility is the stated goal. Respected researchers have tended to choose a standard accent as the target of PI, since these accents are often rated more highly for intelligibility/comprehensibility. Finally, it has been pointed out that, for many learners, for many different reasons, nativeness remains the target and it is these students who are the focus of this research. The following chapter looks at the factors which influence attainment in L2 pronunciation.

3 FACTORS INFLUENCING ULTIMATE ATTAINMENT IN L2 PRONUNCIATION

“... no single factor guarantees a native-like accent
– nor does one single factor, like age, preclude it”

Moyer (2014a, p.19)

L2 , gender, personality, attitude

3.1 AGE OF COMMENCEMENT OF SECOND LANGUAGE LEARNING

In the field of Second Language Acquisition (SLA), it has long been theorised that ‘younger’ means ‘better’. This is the central claim of the *Critical Period Hypothesis* (CPH) initially proposed by Penfield and Roberts (1959) and refined by Lenneberg (1967), which proposes that there is an age-related *window of opportunity* for language learning prior to puberty, after which time, decreasing brain plasticity makes it increasingly difficult to acquire a language and maybe impossible to acquire native-like language competence. In its basic form, the theory focuses on first language acquisition to explain language acquisition in those deprived of language exposure in their early years, such as feral or abused children (see Long, 1990), but it has also been extremely influential in SLA theory. There is considerable agreement, both anecdotally and from research evidence (see, for example, Abrahamsson & Hyltenstam: 2009; Flege & Fletcher, 1992; Flege et al., 1995) to support the notion that the age of commencement of L2 learning, referred to variously as *Age of Acquisition* (AoA), *Age of Onset* (AO) or *Age of Learning* (AoL), is a defining factor in determining the end point of L2 attainment, generally referred to as *ultimate attainment* (UA).

However, there are reasons to challenge a strict interpretation of this model. The CPH posits some sort of age-related neurological change in the developing brain but, as yet, the type of change and its precise location within the brain have not been identified (see Flege et al., *ibid.*; Kinsella & Singleton, 2014; Moyer, 2014b). Furthermore, there is evidence to suggest that neural functions remain plastic well into adulthood (Dick et al. 2005; Herschensohn, 2007). Indeed, Singleton and Lesniewska (2012) argue that, “any decline in L2-learning capacity with age is continuous and linear, which is not in keeping with the usual understanding of the notion of critical period” (p.102).

Challenges to a strict interpretation of the CPH have led to suggestions of a weaker version of age-related decline, such as the *linear increase* hypothesis (Oyama, 1976) learner environment. Both of these propose acquisition, which has been reported in other research (see, for example, Bialystok & Hakuta, 1999; Birdsong, 2006; Singleton & Lesniewska, *ibid.*). This linear decline has been variously defined as starting around 8-10 years (Johnson & Newport, 1989), at 5-6 years of age (Birdsong, *ibid.*; Bley-Vroman, 1989; Long, *ibid.*) or even as young as 12 months for phonology (Ruben, 1997). In a critical review of age-related research into UA, Muñoz and Singleton (2011) use this disparity to sound a note of alarm:

“... if there were clear evidence for the closing of a window of opportunity for language acquisition, surely it ought to be possible for researchers to agree where it is situated. The fact that there is so much disagreement about this matter casts severe doubt on the whole idea of a critical period for language” (p.9).

They suggest that the linear nature of the age-related decline in UA is “more in keeping with what we know about the contour of general cognitive deterioration” (p.11), rather than being a domain specific decline in language learning ability (see also Birdsong, *ibid.*).

A possible explanation for the lack of clarity regarding the timing of the decline in language attainment is that it does not affect all areas of SLA equally. Long (1990), for example, puts phonological maturational constraints at 5/6 years old, but for morphology and syntax at around 15, while Veríssimo et al. (2018) add that “...whereas some subsystems of language acquisition may exhibit sensitive periods that start “closing” in early childhood, other subsystems may remain fully operational until later in life or even throughout the life span” (p.316).

Research seems to suggest that age constraints might affect phonology earlier than other areas of language (Flege et al., 1999; Granena & Long 2013; Huang, 2014). Flege et al. (1999), describe a study involving 240 L1 Korean learners of English whose AoA in the United States varied from 1-23 years. Their English pronunciation was evaluated using a delayed sentence repetition paradigm, which was rated by 10 NSs of American English for degree of foreign accent on a 9-point scale from ‘very strong foreign accent’ (1) to ‘no accent’ (9). Morphology and syntax were evaluated using a grammaticality judgement test. The findings suggested that, as the AoA of the participants increased, their foreign accent grew stronger. Indeed, there were no participants with AoA of 9+ who were judged to have a native speaker accent, supporting the notion of a sensitive period for pronunciation ‘closing’ prior to the age of 9. However, on the grammaticality judgement test, the authors found that AoA was less influential in predicting morphosyntactic attainment, with the superior results for syntax appearing to be influenced by level of education, while lexical scores seemed to be related to the quantity of

English usage by individual participants. Similar findings were obtained by Huang (*ibid.*) in a study which compared grammaticality judgement and perceived nativeness in English of 118 L1 Mandarin speakers whose AoA in the United States was between 5-27 years. Participants underwent a speech perception task, grammaticality judgement task, a speech production task and an oral language proficiency task and AoA was found to be the most influential predictor of all L2 outcomes. However, as with the previous study, the effect of AoA seemed to vary across language subsystems with the author concluding that “the age of learning variable had a stronger impact on speech production than on grammar” (p.397). So, it would appear that there are chronological difference in AoA effects for different language subsystems and that phonology may be one of the first subsystems to be affected by age constraints. Indeed, Moyer (2014b) notes that “L2 learners beyond the age of about 9 or 10 years can exhibit native-like behavior in certain subdomains of language ... but it is particularly unusual to reach a level of pronunciation ability that qualifies as consistently “native” or “near native” (p.444).

However, neither the CPH nor the sensitive period/linear increase hypotheses satisfactorily explain anomalies in phonological attainment. There are cases in the literature of young learners who are judged to have a ‘foreign accent’. In their study of 19 Finnish-born L1 Turkish children, Uzal et al. (2015) distinguish between AoA, which they take to mean the date on which the learner arrived in Finland and Age of Onset (AO), the age at which they were first exposed to the L2 (Finnish). They point out that, while the AoA was 0 for all participants, AO was between 1.7 and 4 years of age. The 19 children, aged between 7 and 17 at the time of the survey, were evaluated by 61 Finnish NS raters. Using a sentence repetition task, raters evaluated the tokens using a 9-point scale from 1 (‘very strong foreign accent’) to 9 (‘no foreign accent’) and judged that only 4 of the 19 spoke without a foreign accent in Finnish. While AO was the most significant factor in predicting accentedness, quantity of L1 home use was almost as significant, with overall quantity of L1 and L2 use also cited as predictors. This led the authors to state that “our findings show that L2 phonological acquisition is affected by multiple factors and that AO alone does not guarantee a native L2 accent” (p.498). Similar findings of early L2 learners who retain an NNS accent are reported elsewhere in the literature (see, for example, Flege et al.,1997; Flege et al., 2002). In the Flege et al., (1997) study, participants were 40 L1 Italian learners of English who had moved to Canada at between 2.6 and 9.6 years old, and who were divided into 2 groups roughly matched for age and AoA, but who differed in their use of Italian: the high use L1 group, who reported using Italian an average of 36% of the time and the low use L1 group, who reported using Italian for an average of 3% of the time. The two groups were rated for accentedness by 24 English NSs using a sentence repetition task and it was found that the high use group were identified as having significantly stronger foreign accents in English, than the low use group. The influence of the quantity, and the quality of exposure to and use of the L2 is dealt with in more detail in Section 2.3

below, but the studies cited above suggest that the quantity of continued L1 use is a factor in determining L2 pronunciation attainment even in those with an AoA of 9 years old and younger.

Furthermore, while relatively rare, there are cases in the literature of late learners who attain a level that is comparable to that of a native speaker, so-called *exceptional learners* (Birdsong, 2007; Bongaerts et al., 1995; Dollman, 2020; Ioup et al. 1994; Moyer, 1999). The study by Ioup et al. (ibid.) compared the language abilities of 2 L1 English speakers, one of whom started learning Arabic at age 21, while the other is described as starting to study Arabic during her final year of an undergraduate course. Both were living in Cairo, at the time of the study, which compared their performance in Arabic with that of NSs of Arabic. Extensive testing in grammaticality judgements, translation, speech production and accent recognition tests found similar results for both groups and led the authors to suggest that the CPH was not an insurmountable barrier to L2 acquisition. The Birdsong (ibid.) study involved 22 L1 English learners of French, whose AoA (defined as age of immersion in the French language, rather than first exposure) was between 18 and 61 years and all of whom were university educated and had resided in the Paris area for at least 5 years. Their production was compared to that of a control group of 17 NS French speakers for voice onset time (VOT) and vowel duration analysed using Signalyze 3.12. VOT refers to the length of time between the release of a stop consonant and the onset of vocal cord vibrations (voicing). Furthermore, 2 read-aloud samples collected from each of the participants were rated for nativeness by 3 NS French raters. Of the 22 NNSs, 2 were found to perform at the same level as the NSs for all 3 tasks, leading the author to hypothesise that “roughly 10% of long-resident, educated Anglophone late learners of French will be able to sound like native” (p.114) at least under the conditions of this study. Similarly, Dollman et al. (2020) in a large-scale survey of 1,843 immigrants to Germany divided participants into 4 groups depending on AoA: under 4, 4-9, 10-16 and over 16, who were rated for strength of foreign accent by German NSs using a 9-point scale for both read aloud and spontaneous speech tokens. The findings indicated that, while there was evidence for the CPH with phonological attainment dropping off at around the age of 10, there were also participants who were able to overcome a late start. Again, high levels of exposure to the L2 was suggested as a factor which helped to compensate for the late start.

Abrahamsson and Hyltenstam (2009) suggest that the apparent existence of exceptional learners may be due to a lack of thoroughness in assessing nativeness. For example, Moyer (1999) conducted a study of 24 high level American learners of German, described as having little exposure to the L2 until after the CPH, which she situates at the age of 10. 4 German NSs rated the learners for nativeness under four different conditions: word-list reading, sentence reading, paragraph reading and free speech

production. While the first three conditions elicited a considerable number of nativeness judgements from the raters, only one participant was judged as native in the free speech condition, suggesting that the type of speech token elicited affected the results. In their own study, Abrahamsson and Hyltenstam (ibid.) studied 41 Spanish/Swedish bilinguals, with an AoA of between 1 and 19 years, who had been identified as native by native speakers of Swedish. The authors carried out what they described as “a battery of 10 highly complex, cognitively demanding tasks and detailed measurements of linguistic performance, representation, and processing” (p.249). The results found that none of the 10 participants over the age of 12 (AoA 13-19 years) and only 3 of the younger learners (AoA 1-11 years) were judged as performing at the same level as the 15 native speakers in the study. The results led them to conclude that, “nativelike ultimate attainment of a second language is, in principle, never attained by adult learners and, furthermore, is much less common among child learners than has previously been assumed” (p.250). In other words, while an L2 speaker might be classed as native or near-native in certain situations or on certain tests, it is very rare that they are classed as such under all conditions. This claim is supported by other research showing that even early bilinguals can be seen to differ from monolinguals (Piske et al., 2001; Hyltenstam & Abrahamsson, 2000).

While Abrahamsson and Hyltenstam (2009) are right to sound the alarm regarding the rigorousness of research into exceptional learners, it is important to note that the two studies cited above, Ioup et al.(ibid.) and Dollman et al. (ibid.) *did* use spontaneous speech samples to evaluate the speakers and in the Ioup et al. study, a variety of different productive and receptive tests were used (as detailed above) to evaluate the L2 competence of participants. Despite the thoroughness of the testing, the two studies still identified some of the late learners as being judged to have NS pronunciation.

Selinker (1972) suggested that exceptional learners who demonstrated what he described as ‘absolute success’ (p.212), who he estimated at approximately 5% of the population, were so unusual that they may be “safely ignored” (p.212) in SLA research. Clearly, 5% of the population cannot be so easily ignored and what is more, as Long (1990) points out, even if there were only evidence of a single late learner achieving native-like competence, this would be sufficient to reject the CPH. These outliers point to the possibility that there might be more to the acquisition of a native-like accent than merely age-related factors.

This is also the view of Flege and Bohn (2021) who, in the *Revised Speech Learning Model* (SLM-r) propose that age-effects in language learning are not the result of a change in how sounds are learned as proposed by the CPH. They argue that factors such as evidence of continuing neural plasticity in later life and the existence of exceptional learners show that humans do not lose the capacity to learn speech at a certain stage of development. Instead, they propose that “all processes and mechanisms used to

develop L1 phonetic categories remain intact and accessible for L2 learning” (p.65) and that “full access” (Flege, 2005) to these processes and mechanisms is available to learners of any age. According to the SLM-r, the age-related differences in acquisition result from the fact that the L2 learner already has an established L1 sound system, so that the developing L2 sound system is shaped by the existing L1 system. As such, it is “virtually impossible for L2 learners to produce and perceive an L2 sound exactly like mature monolingual native speakers of the target L2” (p.25) because the NNS will always be perceiving the L2 through the ‘lens’ of their L1.

In this chapter it has been argued that, while there is evidence of a drop in phonological attainment starting as young as 5/6 years, neurological evidence, plus the existence of foreign-accented early learners and of exceptional learners seems to question the idea of a ‘window of opportunity’ after which the acquisition of NS levels of phonological attainment is impossible. The SLM-r instead proposes that the apparent drop in attainment is the result of interference from an already established L1 phonological system and it is this interference which is the topic of the next section.

3.2 FIRST LANGUAGE INTERFERENCE

The very existence of discrete, identifiable foreign accents in English spoken as a Second Language, such as a French accent or a Japanese accent, provide compelling evidence for the existence of phonological interference from the L1 in L2 pronunciation. In accents languages which share phonemes, this might result in positive interference, helping the learner with L2 pronunciation acquisition. For example, the existence of the voiceless glottal fricative /h/ in German means that L1 German speakers are unlikely to have problems with the articulation of /h/ in words like *hospital* /'hɒspɪtl/ or *harvest* /'hɑ:vɪst/ in British English. However, L1 speakers of Romance languages, such as French and European Portuguese, have been shown to omit the /h/ phoneme in L2 English, and it has been hypothesised that this is because the phoneme does not exist in their L1 phoneme inventory (see John & Cardoso, 2009 - French; Sousa, 2017 - European Portuguese).

Flege et al. (1995) suggest that part of the problem may be that L2 learners struggle at a motoric level to articulate L2 sounds not present in the L1 inventory; in other words, they may lose the motor ability to articulate phones which do not appear in their own language after a certain age (see also Scovel, 1988; Fraser, 2006). This might be particularly true of sounds which are physically complex to articulate, such as the English /θ/ or /ð/ phonemes. Indeed, there is evidence that complex articulations

are acquired later by child L1 learners (e.g., Dinnsen et al. 1990; Watts & Rose 2020). Thus, it might be assumed that learners with L1/L2 pairs having considerable phonemic overlap might be more ‘successful’ in phonological attainment than those whose L1 and L2 have little overlap. However, research seems to suggest the contrary: that learners *are* able to articulate completely ‘new’ sounds that do not appear in their L1, but struggle with ‘similar’ sounds (Flege, 1988; Best & McRoberts, 1989). In his *Speech Learning Model (SLM)*, an earlier version of the SLM-r mentioned in Section 2.1 above, Flege (1995) proposes that, when the sounds are similar, the L2 learner may map the L2 sound onto a similar, already established L1 phonetic category and may fail to distinguish between the two categories. Young-Scholten (2013) gives the following example: “an English speaker learning French will confuse the French /u/ in *vous* ‘you’ with the front-rounded vowel /y/ in *vu* ‘see’ because English only has back rounded vowels and will classify both sounds as the English phoneme /u:/” (p.210).

The SLM sees this not just as an articulatory problem, but also one of perception. It is possible that L1 speakers are not attending to similar sounds in the L2 because the L1 is limiting what they pay attention to in the L2 (see Flege et al. 1995; Hansen, 2004). In the SLM-r, Flege and Bohn (2021) reject the CPH by saying that learner do not lose the capacity to acquire speech as they age, but rather, they learn differently, since the learning is inevitably filtered through existing knowledge, so that “L2 speech learning is profoundly shaped by perceptual biases induced by the L1 phonetic system” (24). That NNS perception is different from that of NSs is supported by research into the judgement of accentedness. For example, Wilkerson (2013) used a rater questionnaire to compare NS and NNS responses to German language speech tokens. While she found that experienced NNS raters performed similarly to NS raters in some areas such as judgement of accentedness and vowel production, in other areas such as intonation/rhythm and comprehensibility their judgements varied, which led her to conclude that, “a reliance on prosodic features alone, such as *speech rate* and *intonation*, has the potential to fool lesser-informed raters into misjudging a NNS for a NS” (p.114). Similar findings by Riney et al. (2005) suggest that NNSs are not perceiving the L2 speech in the same way as an NS. However, it is also fair to mention that other studies have found little difference between NS and NNS judges (Jenkins, 2007; Derwing & Munro, 2009) raising the possibility that NNS judges ‘hear’ the sounds just as NSs do and that perception is not impaired

The Best’s (1995) Perceptual Assimilation Model of L2 (PAM-L2) expands on this idea by suggesting that, as learners gain more experience and exposure to the L2, they may be able to expand existing categories to include the novel L2 sounds and, in some cases, may create new phonemic categories for the L2, which do not exist in the L1. ,show considerable variation,The Second Language Linguistic Perception Model (L2LP) (Escudero, 2005) expands on these models by providing a detailed account of how learners might dynamically adjust their perceptual systems with exposure. From an ‘initial

state’ of learning, when L2 sounds that are similar to L1 sounds are often confused, to a ‘developmental state’ when, with continued exposure and practice, learners begin to differentiate L2 sounds more accurately until the ‘end state’, when successful learners achieve near-native perceptual abilities, effectively distinguishing and producing L2 sounds. Furthermore, the L2LP proposes that it is not only the degree of similarity between L1 and L2 sounds and the amount of L2 exposure which influence L2 perceptual attainment, but also individual difference (ID) variables, such as cognitive abilities (addressed in Section 2.4 below) and motivation (addressed in Section 2.7 below).

That ID variables mediate L1 influence on attainment is also the conclusion of Munro (2018) who, in a synthesis of empirical research in this field, highlights the fact that a reliance on contrastive analysis between the L1 and L2 may be misleading in accounting for pronunciation attainment since it fails “to account for why one learner produces a particular phone correctly and another from the same L1 background does not” (p.278). Possibly, this is partly due to a failure to take regional differences into account. For example, the German voiceless velar fricative /x/ as in *Buch* /bu:x/ (book) is not present in standard British English, but is present in several British regional accents such as Scottish (Cao & Jin, 2017) and Liverpool (Honeybone & Watson, 2006). There appears to be a lack of research into this area, but it would seem reasonable to assume that L1 speakers of these regional accents would not have difficulty with the perception and production of this L2 phoneme, while speakers of standard British English may. This being the case, it is important to be aware that L1 speakers of a given language are not a homogenous group but may vary considerably in their phoneme inventory and phoneme exposure, among other variables. While we can generalise that a speaker of a given L1 might struggle to perceive and articulate certain sounds, this is by no means always the case.

It is also possible that the L1 interference is working, not just at the level of perception and production, but also at the level of orthography. Bassetti (2008) notes the following:

“while children acquiring L1 phonology are only exposed to acoustic input, L2 learners can be exposed to large amounts of orthographic input, from the very early stages of acquisition, after having learnt to read and write another language” (p.10).

Since most late L2 learners are already able to read and write at AoA, they are therefore exposed to both acoustic and orthographic input in the L2. L2 phonological representations might be shaped by the spelling of a word, which is filtered through L1 orthography-phonology correspondences. There is research evidence to support this proposal and, as with phonological interference, this orthographic interference would appear to be both positive and negative. Escudero et al. (2008) found that native Dutch speakers of English, who typically have problems distinguishing /æ/ from /ɛ/, were better able to remember the difference between words containing these two sounds when they were presented with

both auditory and written stimuli, than when they were just presented with auditory stimuli. Similarly, Showalter and Hayes-Harb (2015) showed that native English speakers were better able to remember tones in Mandarin-like words, when spellings which included tone marks were provided, as opposed to spelled words without the tone marks. Thus, it would seem reasonable to conclude that orthographic representations can support the perception and production of L2 sounds.

However, there is evidence that orthography can also be a hinderance. Hayes-Harb et al. (2018) proposed that this might be the case when the L1 and L2 do not share the same conventions. They noted that in German, final devoicing occurs so that, for example, *Rad* (wheel) and *Rat* (advice) are both pronounced with a final [t], but this would not be the case in English. They taught English NSs a series of German-like words which ended in voiced final obstruents pronounced by a native German speaker. One group merely heard the verbal form, while the other group were also shown the written form. Participants of the two groups were then asked to repeat the newly learned words and it was established that those who had seen the written version were more likely to mispronounce the words by producing voiced final obstruents, even though they had never heard the words pronounced in that way. The authors concluded that, “access to the written forms in the input interfered with their acquisition” and that, “these findings indicate a powerful influence of orthographic input on second language lexical-phonological development ...” (p.547). Similar findings have been noted elsewhere in the literature, for example, Piske et al. (2002) and Vokic (2011).

The Orthographic Depth Hypothesis (Katz & Feldman, 1983) notes that writing systems vary in their degree of letter-sound correspondence. In a *transparent* orthographic system (one with close grapheme-phoneme correspondence), the reader is able to ascertain the pronunciation of a word from its written form with accuracy, while in an *opaque* orthographic system (where the grapheme-phoneme correspondence is inconsistent) this method becomes increasingly unreliable. The degree of transparency of a given writing system (its *orthographic depth*) might influence the strategies adopted to decode it. A learner whose L1 is transparent might use grapheme to phoneme decoding, while a learner whose L1 is more opaque might decode at the level of syllables or words, rather than letters. For example, grapheme-phoneme mapping for the English word *thought* might be ineffective, so the word needs to be decoded above the level of the grapheme. Grabe and Stoller (2011) note that some writing systems (e.g. Finnish and Turkish) are almost completely transparent, while others (e.g. English) are more opaque and yet others, such as those with consonantal alphabets (e.g. Arabic) or non-alphabetic scripts (e.g. Japanese) are yet more opaque. It may be the case that learners whose L1 is transparent may have difficulty decoding an L2 whose orthography is more opaque. This was the finding of Georgiou (2021), who noted that L1 Greek speakers, whose L1 orthography is transparent, shown AXB vowel discrimination tests in English (orthographically opaque) were more accurate on

auditory than orthographic tests, suggesting that the learners were being misled by the orthographic input. The author concluded that “speakers of languages with transparent orthographies such as Greek may rely on the grapheme-phoneme correspondence to decode orthographic representations of sounds coming from languages with an opaque orthographic system such as English” (p.118).

To conclude, L1 interference may be working at the level of articulation, speech perception, orthographic interference or a mixture of these variables. Furthermore, L1 interference in perception may be modified by exposure to the L2 and it is this issue which is addressed in the following section.

3.3 L2 EXPOSURE and USE

One would assume that the greater the period of time which the learner has spent acquiring an L2, often referred to as *length of exposure* (LoE) or less frequently *length of immersion*, the more successful they will be. However, the term ‘LoE’ needs to be treated with care as it says nothing about the actual quantity of exposure, the *cumulative exposure*. For example, a distinction is often made between English as a Foreign Language (EFL) contexts, where the learner is not learning in an English-speaking country, and English as a Second Language (ESL) contexts, where the learner is living in an English-speaking country. Clearly, the input offered by an hourly language lesson once a week for a month in an EFL context is in no way comparable to the constant exposure to the L2 which a learner might experience when studying or working for a month in a country where the target language is spoken. It would seem reasonable, therefore, to consider not just the nebulous ‘length’ of exposure, but rather to look at the ‘quantity’ of exposure to and use of the target language, when discussing language acquisition.

That quantity of exposure to L2 affects pronunciation outcomes is supported by much research in the field (Flege & Fletcher, 1992; Flege et al., 1999; Purcell & Suter, 1980; Uzal et al., 2015). In the study by Uzal et al. (2015), mentioned in Section 2.1 above, the authors concluded that AO was the main determiner of perceived accent, but added that quantity of L2 exposure and use was also a significant factor, with those reporting greater exposure to Finnish showing more native-like pronunciation outcomes. It would appear that the effect of exposure on pronunciation is stronger in the initial stages of learning and drops off as the learner becomes more competent. Piske et al. (2001), in a review of relevant research, state that “for highly experienced subjects, additional years of experience in the L2 appear to be unlikely to lead to a significant decrease in degree of L2 foreign accent” (p.199). It seems that, after a certain point in L2 learning, the L2 accent becomes impervious to further change, a phenomenon described in the literature as *fossilization* (Selinker, 1972).

As previously mentioned in Section 2.1 above, quantity of exposure might relate not just to exposure to the L2, but also the quantity of continued L1 use. In a study of bilingual Italian-English speakers living in Canada, Piske et al. (2001) used a delayed repetition technique to compare attainment in those with early vs. late AoA. Late learners were more likely to be perceived as having a foreign accent by NS judges, but the researchers also found that “native Italian subjects’ amount of self-reported use of Italian, ... was an independent predictor of degree of L2 foreign accent” (p.210). Those who continued to speak their L1 frequently tended to have stronger foreign accents than those who spoke Italian more rarely and similar findings have been reported elsewhere (Thompson, 1991.; Flege et al., 1997). However, Piske et al. (ibid.) offer another possible way that the quality, or lack of it, is influencing attainment, when they comment that those in the study who used Italian frequently, “might have heard more frequent tokens of Italian-accented English vowels and consonants than the subjects who used Italian only infrequently” (p.210). If learners are being exposed to heavily accented speech in the L2, then it may be this accent which is their primary model and the ‘developmental state’ predicted by the L2LP (Escudero, 2015), discussed above, is not being triggered, or is triggered to emulate the heavily accented speech.

On the subject of quality of L2 input, it has often been noted that young learners, schooled in the L2, have frequent contact with the L2 in meaningful interactions in various contexts. This sort of meaningful interaction is sometimes unavailable to older learners (Hulstijn, 2015; Krashen, 1985; Moyer, 2011; Moyer, 2015; Muñoz & Singleton, 2011). In ESL contexts, authentic language interactions may be too challenging for the learner, while in EFL contexts, exposure might be scaffolded through simplified, often stilted, teaching materials which may not replicate real-world interactions in the L2. It is possible that relevant and frequent input is more available to school-aged learners, because the school context gives them multiple opportunities to interact in the L2 (quantity), and possibly input that is both authentic and appropriate for their current level of language competence (quality). This has even led some to suggest that the perceived facility with which children schooled in the L2 seem to acquire the language may be more a question of the quality of appropriate input which they are exposed to, rather than age-related factors (Hulstijn, ibid.).

Indeed, reporting on two studies of late learners of English with L1 Dutch, who were deemed highly successful in their acquisition of a native-like accent, Bongaerts et al. (1997) state that the highly successful learners in their studies were “consistently exposed to a large amount of input from native speakers of the target language” (p.461). Similarly, referring to extensive work in the area, Moyer (2009) states that “if quality and quantity of input are limited ... attainment predictably suffers” (p.169). The suggestion here is that the impediments to late-learner achievement proposed by the CPH can be overcome if sufficient, good quality exposure to the L2 is provided. In fact, Muñoz (2014), in a

study of 160 L1 Spanish learners of English, suggested that input might be more important than AoA in determining oral performance outcomes. She used self-report questionnaires to compare the relative influence of AoA and input on oral performance of the learners whose AoA was between 3 and 15.5 years. Oral performance was measured in terms of speech rate, lexical range, syntactic complexity and accuracy, rather than measures of accentedness, but the findings indicated that “input has a stronger association with measures of oral performance than starting age” and that “cumulative exposure and, above all, contact with high-quality input are good predictors of learners’ oral performance in the foreign language” (p.1).

That exposure to and use of the L2 is fundamental to outcomes is also borne out in research into exceptional learners, learners who achieve native-like competence discussed in section 2.1 above. Self-reporting studies of exceptional learners suggest that they seek out such meaningful interactions in the target language and deliberately avoid speaking their L1 (Moyer, 1999; Piller, 2002). It would seem that, consciously or unconsciously, these exceptional learners are aware that it is both the quantity and the quality of their interactions in the target language which will help them achieve their learning goals.

So far, phonological attainment has been explained largely in terms of global experiential factors which might be expected to affect all learners to a greater or lesser extent, namely AO, L1 interference and quantity/quality of L2 exposure. However, ID variables, such as cognitive and psychological differences between learners which might facilitate or hinder the attainment of pronunciation goals, also need to be taken into consideration. Mora and Darcy (2019) suggest that “pronunciation is the area of language with the largest individual variation in performance, compared to grammar or vocabulary” (p.95) and research has shown that these differences remain significant, even after experiential factors are controlled (Bradlow et al. 1997; Golestani & Zatorre, 2009). In the following sections, various ID variables which may impact phonological attainment, namely language aptitude, gender, personality, attitude and identity are discussed.

3.4 LANGUAGE APTITUDE

It has often been thought that some people are somehow ‘better’ at learning languages than others, and this perceived language facility is referred to as *language aptitude*. This sort of aptitude for languages is described by Abrahamsson and Hyltensam (2008) as “a largely innate, relatively fixed talent for learning language” (p.485) while Wen and Skehan (2021) define it thus:

“foreign language aptitude generally refers to a set of specific cognitive abilities that enable some individuals to learn additional languages more quickly and efficiently than their peers when all other language learning success factors (e.g. time, quality of instruction, motivation) are equal” (p.6).

It has been variously reported that anywhere between 5 and 15% of learners might possess this aptitude (Birdsong, 1999; Selinker, *ibid.*; Seliger et al.1975). Furthermore, Abrahamsson and Hyltenstam (*ibid.*), in a review of recent studies into aptitude, observe that “child learners seem to attain a native-like command of the L2 regardless of high or low aptitude” (p.481), which seems to mean that, while aptitude may be a factor in late-learner L2 attainment, it does not appear to be implicated in early L2 acquisition.

Trofimovic et al. (2015), in an overview of factors affecting L2 pronunciation development, define language aptitude as encompassing the following areas: short-term memory, phonetic coding, grammatical sensitivity, rote learning (ability to form sound-meaning associations), mimicry, inductive learning (ability to infer rules from linguistic information) and musical ability. Various tests have been established to measure language aptitude. Among the most widely used is the Modern Languages Aptitude Test (MLAT), proposed by Carrol and Sapon (1959) in which the focus is on perception, not production and which aims to assess language aptitude in general rather than phonological skills. In the subsequent years, MLAT has been added to and adapted (see, for example, the High-Level Language Aptitude Battery, (Hi-LAB), Dougrty et al., 2007; LLAMA aptitude tests, Meara, 2005) and new variations of aptitude tests have been developed (see Skehan et al., 2016 for an overview). However, to date, there would appear to be no academically recognised test which claims to focus on pronunciation aptitude, in other words, language aptitude as it relates specifically to pronunciation attainment. Indeed, in a comprehensive analysis of 66 studies of language aptitude, Li (2016) noted that language aptitude tests were often flawed since, “different aptitude components demonstrated differential predictive validity for different aspects of learning” (p.801). In the absence of an agreed definition of pronunciation aptitude, this section looks at research into various language aptitude variables mentioned by Trofimovic et al. (*ibid.*) above, which have been specifically linked to pronunciation namely, short-term memory (which is expanded to include other aspects of cognitive control linked to pronunciation – inhibitory control, attention control and phonological self-awareness), auditory processing skills, mimicry ability, and musical aptitude.

Furthermore, the definitions of aptitude given above see it as a stable trait, one which does not change to any great extent throughout the lifespan, However, there is a growing body of research which suggests that some of the traits assigned to language aptitude in general, and pronunciation aptitude in

particular, may be amenable to change, such as working memory (Biedroń & Skehan, 2017; Singleton, 2017), phonological self-awareness (Meritan & Mroz, 2019; Trofimovic et al., 2016) and auditory processing skills (Kachlicka et al., 2019; Saito, 2023). The remainder of this section looks at research into each of these aptitude variables related to phonological attainment and their responsiveness to intervention.

3.4.1 Cognitive Control

Mora and Darcy (2017) state that cognitive control “refers to the general-purpose control mechanisms that regulate information processing skills and behaviour” (p.96). As relates to L2 phonological acquisition, Mora and Darcy (2023) posit working memory, inhibitory control and attention control as the principal aspects of cognitive control. The following section looks at these 3 dimensions of cognitive functioning as they relate to pronunciation attainment and adds 2 further dimensions: auditory processing and phonological self-awareness which have also been linked to attainment in pronunciation.

2.4.1.1 Phonological Short-Term Memory /Phonological Working Memory

In the field of L2 pronunciation, the feature of working memory relevant to aptitude is usually referred to as Phonological Short-Term Memory (PTSM), which Mora and Darcy (2017) define as “a short-term phonological store for verbal information which allows individuals to encode phonological units and their sequential order in the form of auditory traces that can be sustained in memory” (p.98). While PTSM refers to memory alone, Phonological Working Memory (PWM), describes the ability “to simultaneously manipulate this information for completing some cognitive tasks” (Wen et al., 2019, p.365), so it includes the notion of manipulation, as well as storage of auditory information and, as such, PTSM can be viewed as a sub-component of PWM (Singleton, 2017). PTSM and PWM capacity are often assessed using the forward and backward digit span tests (Darcy et al., 2015; Grivol & Hage, 2011). In the forward digit span test, participants are presented with a sequence of numbers and are asked to repeat them in the same order so as to measure the ability to temporarily store and recall information, PTSM. The maximum sequence of numbers which can be recalled in the correct order defines the participants forward digit span. The backward digit span test involves a similar process, but participants are required to repeat the numbers in reverse order, thus testing not just PTSM, but also PWM. Other tests include non-word repetition and recognition (Kormos & Sáfár, 2008; Safronova & Mora, 2013). In the former, participants are presented with a series of non-words,

that is, meaningless words which follow phonotactic rules of the language being tested, and which participants are asked to repeat exactly as heard to test their PSTM. The recognition task builds on the repetition task, by requiring participants to manipulate the non-words, such as identifying whether two non-words are the same or picking a non-word from a set after hearing it once, thus testing PWM.

Miyake and Friedman (1998), in a comprehensive review of ID variables in L2 proficiency make the claim that “working memory (WM) for language may be one (if not the) central component of this language aptitude” (p.339), but it is unclear to what extent this can also be claimed for phonological attainment. Trude and Tokowicz (2011) report on a study of L1 English speakers with no previous experience of Portuguese, half of whom had some knowledge of Spanish. After a training session giving them a brief introduction to Portuguese phonetics, participants were assessed for working memory, using an operation span task where participants were shown a mathematical operation and asked to judge if the presented answer was correct. After pressing a key to indicate their decision, they were shown a Portuguese word. This procedure was repeated between 2-6 times and then participants were asked to write as many of the words as they could remember. Their recall of the Portuguese letter-to-sound correspondences was then tested by asking them to read out Portuguese words which had not appeared in the training phase of the study. They found that “individuals with higher working memory performed more accurately than individuals with lower working memory” (p.274). As the working memory they were referring to was specifically related to phonological recall, their research supports the idea that PWM aids phonological attainment.

However, it is also important to note that there are studies which have not found a relationship between the two variables. For example, Hu et al. (2013) used a read-aloud task to assess pronunciation attainment, a digit-span and a non-word recognition task to assess working memory and a self-report questionnaire to assess behavioural features of 109 L1 German learners of English and found that PWM was not a predictor of L2 pronunciation ability in advanced learners. Similarly, Safronova and Mora (2013) tested the role of acoustic memory (AM) and PSTM in the perception of the English /i:/ /ɪ/ distinction in Spanish/Catalan intermediate to advanced L2 learners and noted that “participants’ L2 vowel perception was not significantly affected by individual differences in AM and PSTM” (p.384). It could be, therefore, as suggested by Tsai et al. (2016) that PWM has more of an influence in the early stages of L2 acquisition and plays less of a role as learners gain proficiency. However, other studies (see Biedroń, A., & Szczepaniak, 2012; Perani, 2005) suggest the exact opposite, that there are indications that working memory is more fully engaged at the advanced levels of L2

acquisition. Part of the discrepancy may lie in the fact that, while some of the research cited above is based on the rating of individual sounds and words (Safronova & Mora, *ibid.*; Trude & Tokowicz, *ibid.*), other research relies on read-aloud and sentence repetition tasks (Hu et al., *ibid.*).

It might also be the case that working memory is not a discrete skill that can be isolated from other aspects of cognitive control in research paradigms. For example, Trude and Tokowicz (*ibid.*), in the study of working memory mentioned above, proposed that individuals with higher working memory are better at inhibiting L1 (see also Kane & Engle, 2003). The following section explores the possible role of inhibitory control in pronunciation attainment.

2.4.1.2 Inhibitory Control ✓

Borragan et al. (2018) note the following:

“Multilinguals have to manage different languages to control verbal speech on an everyday basis. They have to select the language that is needed at every specific moment and suppress interference from the situationally irrelevant languages. ... Broadly speaking, inhibitory control refers to the suppression of interfering information or prepotent responses” (p.1)

This is the central idea behind the Green (1988) Inhibitory Control (IC) Model, which suggests that bilingual speakers have the ability to block interference from their other language(s), in order to communicate efficiently.

Providing evidence for inhibition in speech production, Darcy et al. (2014) measured inhibitory control in the L1 by having participants memorise 6 words from 3 categories (vegetables, occupations and animals). They then practised some items in two of the categories. This was expected to cause inhibition of unpractised words within these two categories, while the unpractised category served as a control as there would be no inhibition. Participants were then tested on their reaction times in retrieving practised and unpractised items from the two categories, in an ABX task, where they had to choose if the X token was more similar to the A or the B token in order to judge their ability to inhibit their previous learning. The findings showed that learners with high inhibition scores, performed better on the ABX task than those with lower scores, leading the authors to suggest that “perhaps learners with higher inhibitory skill are able to deactivate (or inhibit) the language not in use more efficiently” (p.127).

The Inhibitory Control (IC) Model mentioned above also proposes that inhibition of the L1 or the dominant language will be more difficult than inhibition of the L2. Levy et al. (2007), related inhibition to the concept of L1 *attrition* – the idea that, after a period of L2 immersion, speakers have difficulty retrieving L1 lexis. After a series of tests requiring L1 English speakers learning Spanish to name Spanish objects, they found an attrition effect in the L1 and proposed that the L1 lexis had been suppressed. Furthermore, they suggested that the greater the asymmetry between L1 and L2 ability, the greater the inhibition for L1 words. Evidence for this second claim was provided by a series of experiments by Costa and Santesteban (2004). The participants were L1 Spanish learners of Catalan and L1 Korean learners of Spanish, who were shown a series of pictures, which they were randomly instructed to name in one of their two languages. They found that, while highly proficient bilinguals were able to switch between languages with equal ease, those with less proficiency in their L2 found it harder to switch from their L2 to their L1 than vice versa. This led the authors to conclude “if the bilingual has not achieved comparable performance levels in her two languages, then lexical selection would make use of inhibitory control to ensure selection in the intended language” (p.505). In other words, the balanced proficient bilinguals did not have to rely on inhibition, while the non-bilinguals had to work harder to suppress the L1 than they did to suppress the L2, leading to the observed slower response times. The more dominant language has to be inhibited to a greater extent than the less dominant language.

Levi-Ari and Peperkamp (2013) relate inhibition to the notion of L1 interference in an L2 discussed in Section 2.2 above, except they focus on the opposite – how an L2 might influence the L1 in late bilinguals. In one experiment with L1 English speakers who were late learners of French of upper-intermediate proficiency, the authors noted that “the lower inhibitory skill late English–French speakers had, the more French-like their pronunciation of word-initial voiceless stops in English was, even though the task was solely in English and the experimenter was a native speaker of English” (p.327). This led the authors to conclude that, “poorer inhibitory skill results in greater co-activation of the competing linguistic units, and that this co-activation can ultimately influence representation.”

Taking all of the above into account, it would appear that the ability to inhibit interference from other languages is an advantage for late learners, but that highly proficient learners and balanced bilinguals may be able to dispense with inhibition, so that it no longer plays a role in language performance.

2.4.1.3 Attention Control ✓

Mora and Darcy (2017) describe attention control as “the cognitive mechanism that enables individuals to flexibly and efficiently switch their focus of attention between tasks or mental sets, in this case linguistically relevant speech dimensions” (p.98).

Darcy et al. (2014) measured attention control and phonological processing in 16 L1 Spanish learners of English and 18 L1-English learners of Spanish. Participants were presented with a series of non-words and, for each one, they were asked about one phonological dimension of the sound: a) was it nasal or not? or b) was it spoken in the first language of the participant or not? Two trials were carried out: a repeat trial where the phonological dimension stayed the same, and the switch trial which alternated between the two dimensions. The authors reported that both groups performed faster on the repeat trial and they used the response time difference between the two as the measure of attention control. L2 phonological processing ability was measured using an ABX discrimination task and L2 pronunciation accuracy measured in a delayed sentence-repetition task, in other words, both a perception and a production task. The attention control and pronunciation scores were then compared. The authors noted that “more efficient attention control (operationalised as lower shift-cost scores) was associated with more accurate performance in ABX (for the L2-English learners)” (p.126). For the L2-English learners, the findings for the production task were less clear and, for the L2-Spanish learners “there was no relationship between attention control and production scores” (p.126). Therefore, there would appear to be some evidence that attention control influences L2 perception, but it is less clear whether this transfers to production. Possibly, in online speaking tasks, attention is focussed primarily on message content rather than on acoustic features of the exchange. A similar study by Safronova and Mora (2012) found a similar advantage for attention control in L2 perceptual learning of L2 sounds, but did not test for production.

Mora and Mora-Plaza (2019) examined three subcomponents of auditory attention – selective attention, attention switching and inhibition in their effects on the training of 17 Catalan-Spanish bilingual undergraduate English learners in their discrimination of two English vowel contrasts /æ/-/ʌ/ and /i/-/ɪ/. Findings suggested a correlation between participants’ perception gains post training and auditory selective attention for an ABX discrimination accuracy test. However, pre-test, there was also a correlation between discrimination accuracy and both auditory attention switching and L1 inhibition scores for the /æ/-/ʌ/ contrast. These results suggest a link between attention and perception which, at least in the area of auditory selective attention, may be amenable to training.

Therefore, there is some evidence that attention control influences L2 perception, but less clear whether this transfers to production. It is also worth noting that each of the research studies above focussed on individual phonemes or words, and it unclear how this would generalise to spontaneous speech production. Clearly, more work needs to be done on the role of attention control in both perception and production of L2 speech sounds and this work needs to extend to its influence on more meaningful stretches of speech.

2.4.1.4 Phonological Self-Awareness

One further aspect of cognitive control which has been linked to phonological attainment is that of phonological self-awareness: the learner's awareness of their own production. Research comes from the field of skill acquisition, where much attention has been paid to the *Dunning Kruger effect* – the oft-observed discrepancy between how people judge their own abilities compared with how others judge them (Kruger & Dunning, 1999; Carter & Dunning, 2008). More recently, this effect has also been observed in L2 pronunciation (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al. 2016). Trofimovich et al.(ibid.) tested 134 L2 speakers of English from different language backgrounds on a picture narrative task, and their own judgements of their nativeness and comprehensibility were compared to the ratings of 3 English NSs. The authors noted that, for many L2 speakers, their self-rating of accent and comprehensibility only weakly dovetailed with that of listeners and that “only about a third of all L2 speakers were fully calibrated with listeners in their self-assessment” (p.133). Similar findings are reported by Lefkowitz and Hedgcock (2002) who, in qualitative data obtained from interviews and questionnaires with 282 L1 English learners of Spanish and French, noted, “a tendency to inflate oral performance self-assessments” (p.223). That learner self-awareness may be implicated in pronunciation attainment is highlighted by many writers in the field (Dlaska & Krekeler 2008; McDonald, 2018; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovic et al., 2016). Saito et al. (ibid.), in a study of 106 Japanese ESL learners, found a correlation between high comprehensibility ratings as rated by 5 NS raters and learner self-assessment of their comprehensibility. Conversely, McDonald (2018) points out that if learners have an unrealistic view of their own accent, then they may have a limited understanding of “which “bits” of it cause problems with intelligibility” (p.253). Meanwhile, Yoshida (2016) offers the following warning of the possible consequences of unrealistic self-evaluation as relates to pronunciation attainment:

“... while it’s not practical to set our goal impossibly high, we also can’t afford to set it too low. It’s not helpful for students to become too complacent and to believe that their pronunciation is fine when, in fact, it may not be easily understood by anyone other than their own teacher and classmates” (p.2).

More evidence that phonological self-awareness is instrumental in promoting positive outcomes in pronunciation attainment comes from interviews with exceptional learners, who frequently identify self-monitoring as a strategy which they employ in their language learning (Moyer 2014a; Moyer, 2018; Muñoz & Singleton, *ibid.*). Moyer (2014a) looked into a series of self-reports of learners who were rated as near-native speakers in their L2 and commented, “nearly every exceptional learner mentioned self-monitoring” (p.8). Moyer (2015a) argues that for ‘a motor-based skill like accent, a strong metacognitive orientation is required’ (p.11). The forms of self-monitoring reported by studies into successful learners are varied. Moyer (*ibid.*) notes that exceptional learners report frequently asking for and incorporating feedback on their production accuracy. This is also a trait noted by Ioup et al. (1994) in their description of the exceptional language learners discussed in Section 2.1 above. Of one of the exceptional learners in their study, they say she reported that “when she received error feedback, she made mental note of the information it provided” (p.92). Other reports suggest that successful learners are often methodical in their approach to learning, strategizing ways to work on accent and to evaluate their own progress (Biedroń & Szczepaniak, 2009; Hişmanoğlu, 2012). Moyer (2021) comments on exceptional learners’ particular attention to difficult phonological features, including practising problem sounds (see also Moyer, 2014a) or rehearsing/imitating sections of speech (Ghyasi et al., 2013).

While difficult to prove conclusively, it would appear that self-monitoring is helping exceptional learners to refine their pronunciation in order to achieve their goals, while less ‘successful’ learners show a disconnect between how they judge their accent and how NSs judge them, suggesting a failure or an inability to self-monitor. The role of PI in improving phonological self-awareness is addressed in Section 3.4.2 below.

2.4.1.5 Summary - Cognitive Control

The following conclusions can be drawn: there appears to be a link between phonological attainment and PWM; that inhibitory control is necessary for late learners, but unnecessary for proficient and bilingual learners; that attention control is influential in L2 perception at the

phoneme and word level, but its link to production is less clear; and that learner ability to self-evaluate pronunciation may correlate with high pronunciation outcomes.

The complexity of the above discussion of the role of cognitive control in pronunciation attainment and the relationship between each aspect of cognitive control may explain why conclusions are somewhat vague. Darcy et al. (2015) compared three cognitive abilities: working memory, attention control and processing speed with L2 phonological acquisition scores and concluded that, “there was not one single cognitive variable with which all three phonological tasks and the overall phonological score would correlate” (p.71). It would appear that individual learner backgrounds, task effects and different aspects of cognitive control may be interacting with each other to cloud potential findings and that other variables may also be at play. For example, in the Hu et al. (ibid.) study mentioned in 2.4.1.1 above, the authors concluded that that, “phonetic coding ability and empathy, but not phonological working memory, predict L2 pronunciation aptitude in advanced learners” (p.366). The following section looks at the role of phonetic coding ability in phonological attainment, while ‘empathy’ is addressed in Section 2.4.3.

2.4.2 Auditory Processing /Phonetic Coding Ability✓

Saito (2023) describes *auditory processing* as “a complex of domain-general perception abilities related to encoding the acoustic characteristics of sounds” (p.2). This refers to the listener’s ability to precisely perceive dimensions of incoming acoustic information such as pitch, resonance (formants), intensity and duration. which, according to Saito, is effectively synonymous with the notion of “having a good ear” (p.2). Shao et al. (2023) expand on this by noting that it includes two constructs: auditory acuity – the ability to hear subtle details of sound (perception) and audio-motor - integration how well this perception can be translated into motor action (production). When referring specifically to the processing of speech sounds, the term *phonetic coding ability* is often used.

Some researchers suggest that auditory processing is at the very heart of L1 acquisition (Müller et al, 2012) and given this believed importance in L1 acquisition, L2 researchers have begun to study the possible influence of auditory processing on L2 acquisition. Kachlicka et al. (2019) studied the English abilities of 40 L1 Polish speakers living in the UK for between 1- 6 years. All had arrived in the UK after the age of 16, had considerable exposure to English and were motivated to change their accent. They were subjected to an extensive array of auditory processing and speech perception tests, along with grammaticality judgement and language aptitude tests. The results suggested that, in these late

learners, “auditory processing was the strongest predictor of L2 learning success, exceeding the predictive power of language input characteristics such as age of acquisition and length of residence, as well as a test of language learning aptitude” (p.20). The authors warn, however, that the directionality of the effect was unclear and that it was possible that it was language learning experience, which was influencing auditory processing ability, rather than the other way around.

Saito (2023) describes this important influence of auditory processing on language outcomes in late learners in the Auditory Precision Hypothesis-L2. Summarising a series of projects carried out by his team involving over 400 L2 English speakers from various language backgrounds, he describes how participants were assessed for comprehension and production ability, carried out behavioural and EEG (electroencephalogram brain scan) auditory processing tests, and were interviewed for language learning background data. The results suggested that “performance scores were EQUALLY associated with biographical and auditory processing factors” (p.7) and this led the author to suggest that “individual differences in auditory processing may serve as a moderate-to-strong determinant of post-pubertal L2 speech acquisition” (p.8).

Interestingly, both the Kachlicka et al. (ibid.) report and the Saito (ibid.) summary cited above propose that some form of auditory training might be effective in addressing auditory shortcomings. Early research has shown positive results in this area. Saito et al. (2022) found that formant discrimination training and formant + phoneme discrimination training with Japanese learners of English “significantly enhanced different areas of participants’ auditory and speech abilities with medium effects” (p.29). In other words, the specific auditory training transferred to more general improvements in auditory speech processing, leading the authors to recommend more research in this area.

It has also been proposed that the *speed* of auditory processing ability may also be a factor in attainment. Golestani et al. (2007) used brain imaging tests to study the relationship between brain structure and the learning of foreign speech sounds and concluded “it is possible that individuals with a particular type of cortical response to speech sounds, which arise from a particular pattern of anatomical connectivity, can learn to hear new speech sounds more quickly” (p.581). Darcy et al. (2015) give a possible explanation for this effect when they hypothesised that “faster processing would free resources for handling incoming speech material” (p.69). In other words, faster processing might avoid overload of the PSTM which, as discussed in Section 2.4.1.1 above, may also be critical for phonological acquisition. The authors tested this hypothesis by assessing the processing speed of 30 L1 Korean speakers using a speeded naming task where participants were asked to describe the size, colour and shape of as many of the displayed items as possible in 30 seconds in their L1. Phonological

processing was measured for segmental categorization, word stress and lexical decision making. The relationship between processing speed and L2 phonological processing scores were found to relate significantly only to the processing of word stress, but not with segmental categorization or lexical decision tasks. There appears to be little research into this specific area and thus, while auditory processing ability has been identified as an aptitude variable in phonological attainment, the specific role of speed of processing remains unclear.

Meanwhile, other research has extended the definition of auditory processing to include, not just perception, but the ability to reproduce the auditory input. Saito, Kachlicka, Sun, and Tierney (2020) describe how 100 L1 Polish late learners (19-45 years) of English were tested on their L2 pronunciation proficiency and this was then compared to AoA, LoE and auditory processing scores for perception and production. They concluded that the research, “has further extended the hypothesis that auditory processing is a crucial determinant of multiple dimensions (perception, production) and phases (rate of learning, ultimate attainment) of L2 speech learning” (p.12). They suggested that this was the same process as that involved in L1 acquisition and that individual differences in auditory perception skills could go some way to explaining variation in attainment seen in late learners. They also noted that, like L1 auditory processing, it appeared to be “negatively associated with chronological age” (p.12).

In an attempt to tease out the relative importance of auditory acuity and audio-motor integration skills, Shao et al. (2023) tested the auditory processing ability of 47 L1 Chinese learners of English. The methodology involved a shadowing task where participants attended two weeks of choral repetition training. All participants were found to have improved comprehensibility and accentedness scores, but the degree of improvement varied widely. Improvements were found to be linked to participant auditory processing scores and, more precisely, that “the results specifically showed that audio-motor integration, rather than auditory acuity, was related to the rate of learning success” (p.23).

To summarise, it would appear that auditory processing ability, both in terms of perception and production, plays an important role in L2 pronunciation acquisition. It would appear to decline with age, but auditory training may go some way to check the decline. However, Shao et al. (ibid.), in the study mentioned above, note that the less successful participants “may not be able to fully take advantage of shadowing training, ... because they struggle to remember acoustic patterns ...” (p.23). Thus it is possible that auditory processing is closely linked to PWM, as discussed in Section 2.4.1.1 above, but the nature of that link is unclear.

Auditory processing, specifically phonetic coding ability, has also been linked to the ability to mimic unfamiliar sound and this is the topic of the following section.

3.4.2 Mimicry Ability

Li (2016) distinguishes between phonetic coding and mimicry in that the former involves comprehension, whereas the latter is specifically related to sound production. Reiterer et al. (2013) describe L2 mimicry ability as a talent which allows learners to accurately imitate foreign speech despite a late AoA and suggest that cognitive factors may be involved, such as working memory, auditory long-term memory, as well as other factors such as musicality and personality. While the reason for individual differences in imitation ability are unknown, their research suggests structural and functional differences in the brain as a part of the explanation (Reiterer et al., 2011).

Mimicry was proposed by Purcell and Suter (1980) as one of the predictors of pronunciation success and has been included as an aptitude variable by researchers in the field of pronunciation acquisition (Hinton, 2012; Piske et al., 2001; Reiterer et al., 2011). Piske et al., (2001), in an overview of factors affecting degree of L2 accent, note that the ability to mimic unfamiliar sounds “has repeatedly been identified as a significant and independent predictor of degree of L2 foreign accent” (p.202). It has also been identified as a feature of exceptional learners (see Moyer, 2013; Moyer, 2015).

While there has been some research suggesting that mimicry ability is not a factor in pronunciation attainment (see for example, Flege et al., 1995; Flege et al., 1999; Suter, 1976), the majority of studies suggest otherwise. Strapp and Center (1999), in a study of the ability of 28 L1 Japanese learners of English to mimic English words containing the /r/ and /l/ phonemes, found that mimicry appeared to be a distinct skill which was unrelated to age. Reiterer et al. (2013) studied 138 German-English bilinguals for their speech imitation ability in Hindi, a language which was unknown to any of the participants. Using functional neuroimaging, they found that high achievers seemed to have a larger *articulation space*, defined as the range of possible movements and positions of the speech articulators involved in the production speech sounds. This led them to conclude that the evidence indicated that individual differences in speech imitation ability play a major role in speech production and go so far as to suggest that “for our high ability speakers we can refute a critical period for phonetic learning and propose instead a successful mastery of native speech where interference from L1 is not at work” (p.11). In a related study, Reiterer et al. (2011) used fMRI brain imaging to compare brain activation in those they had identified as good and poor mimics. They reported that low ability imitators “showed significantly higher amounts of activation” (p.9), which they proposed placed a higher cognitive load on PWM than on the high ability mimics.

Similar links between phonological attainment and mimicry ability have been found in other studies. Hinton (2012) studied a group of 10 Polish students, none of whom had previously studied French, and played a series of French words and short phrases which they were asked to repeat after hearing once. It could, of course, be argued that this was therefore testing PWM, rather than mimicry. The study could also be questioned on validity grounds since the speech samples were judged by only one rater and that rater was English, despite the samples being for French words and phrases. This may partially explain the author's finding that the correlation between production and mimicry ability "was so strong as to be difficult to believe" (p.108). While methodological concerns make the findings questionable, similar results have been achieved elsewhere (Kendrick, 1997).

While the evidence for mimicry playing an important role in pronunciation is strong, Piske et al. (2001) make an important point when they note that it is unclear the extent to which "some people are born with this ability, or whether it develops as a result of certain still-to-be identified experiences" (p.203). Research into the effectiveness of instruction in promoting mimicry ability is addressed in Section 3.4.4 below.

3.4.3 Musical Aptitude

Anecdotally, musical ability has often been linked to language acquisition in general and to phonological acquisition ability in particular, but research findings have been inconclusive in both areas. Abrahamsson and Hyltenstam (2008) comment that "the popular belief that musical and linguistic talent are closely connected has generally not been borne out in empirical research" (p.485), while Moyer (2013) says of the musical talent/phonological acquisition link "it is a popular idea, but there is little evidence to support it (p.55).

The establishment of concrete findings in this area would appear to be hampered by a lack of consistency in defining what is meant by the nebulous term, 'musician' and how musical ability is assessed. Milovanov and Tervaniemi, (2011) warn of the danger of treating 'musicians' as a homogenous group when they note "it is not yet certain how much the specific demands set up by a given musical genre or instrument can modulate the underlying brain functions" (p.1). It might be argued, for example, that there is little overlap in the skills required to play drums in a popular music band and those required of a violinist in a classical orchestra and yet both would be classed as musicians. Indeed, an important theory on the relationship between musical skills and speech development, the OPERA Hypothesis (Patel, 2011), suggests that music training alone may not impact speech processing. Rather, Patel proposes that "the benefits of musical training depend on the

particular acoustic features emphasized in training, the demands that music places on those features in terms of the precision of processing, and the degree of emotional reward, repetition and attention associated with musical activities” (p.12).

That musicians cannot and should not be treated as a homogenous group is borne out by Christener and Reiterer’s 2013 study into the relationship between singing ability and speech imitation ability. They tested 41 L1 German participants for their singing ability as judged by expert raters and their speech imitation ability in a read aloud task of a well-known Aesop fable in English, plus their repetition of sentences in Hindi, a language unknown to any of the participants. They concluded that “singing performance is a better indicator of the ability to imitate speech than the playing of a musical instrument” (p.1), since singers are required to focus on aspects of vocal production, which instrumentalists do not. However, it is important to note that the study only tested singers, so their conclusions about instrumentalists came from other studies which were, possibly, methodologically different.

Many linguistics have overcome the problem of the lack of homogeneity between musicians by determining musical ability using established tests, rather than relying on previous musical experience. For example, Christiner and Reiterer (ibid.) and Coumel et al. (2023) used the AMMA test (Advanced Measures of Music Audiation, Gordon, 1989) to judge level of musical ability. This is a purely perceptual test, which assesses the ability to identify rhythm or tonal differences in passages of music. Milovanov et al. (2009) and Vangehuchten et al. (2015) used the Seashore Test of Musical Ability (STMA), which tests the participants ability to discriminate pitch, loudness, tempo, timbre, and rhythm. Meanwhile others, for example Sleva and Miyake (2006) and Zheng et al. (2022) have used variants of the Wing Measures of Musical Talents (Wing, 1962) to test participants’ skill in perceiving, memorising and reproducing melodic and rhythmic features of a passage of music, thus focussing on both perception and production. Until there is some agreement on which tests linguists should use to establish ‘musical ability’, comparisons between the findings of these various tests are subject to validity issues.

Some studies have avoided this problem by focussing on specific skills, within musical ability, which might facilitate pronunciation acquisition. For example, Lorenzen (2019) asked, “is there a correlation between musical memory and the ability to recognize Mandarin tonemes?” She used the Mandell Tondeaf Test : Mandell, J. (n.d.) to judge participants ability in pitch perception and melodic memory, and found a positive correlation between these abilities and Mandarin toneme recognition. However, the researcher herself warned of the complex nature of interactions between individual differences in

the participants and variables such as musical ability, concluding, “it is important not to oversimplify the effect of any one influence in isolation from the others” (p.316).

Dolman and Spring (2014) tested 34 L1 Japanese learners of L2 English using the STMA test, mentioned above, to assess musical ability. They tried to ascertain if the specific sub-domains of musicality assessed in the test, namely pitch, duration (time), timbre, rhythm, loudness and tonal memory efficacy showed a greater correlation with the ability to produce 6 English phonemes which they predicted would be problematic for L1 Japanese speakers: /θ/ /ð/ /ɾ/, /l/, /f/, and /v/. Participants listened to an English word containing one of the phonemes spoken by a NS and were then asked to produce the word themselves. The tokens were then judged for nativeness on a 5-point scale. They concluded as follows:

“Though overall musical aptitude did not correlate to overall pronunciation ability for Japanese learners of English, a correlation was found between the ability of the learners to pronounce /ɾ/ and /l/, and the time subtest of the Seashore Musical Aptitude test. These results suggest that there are some potential benefits to having musical ability when learning foreign language pronunciation, but does also point to the fact that some caution should be used before assuming that all types of musical ability will help every aspect of learning pronunciation” (p.9).

While the above research is informative about the relationship between musical ability and phonological short-term memory in L2 speech production, it does *not* look at how musical ability might influence production in spontaneous speech. Zheng et al. (ibid.) tried to address this issue by using a picture narration task with L1 Chinese learners of English in order to obtain spontaneous speech tokens. They assessed participants’ general auditory processing using an AXB discrimination task. Musical ability was assessed using a perception task in which participants listened to two musical phrases and decided if they were the same or not, and a production task in which they were required to listen to, remember and replicate (sing or play) musical phrases. They found that, while differences in auditory processing were related to L2 pronunciation proficiency, musical aptitude was only weakly linked to prosodic aspects of pronunciation, which was probably the reason they entitled the paper, “Successful second language pronunciation learning is linked to domain-general auditory processing rather than music aptitude” (p.1).

In conclusion, while the evidence of a clear link between the ambiguous term ‘musical ability’ and pronunciation acquisition is thin, it may be that the two skill sets share certain abilities which are common to both, such as auditory memory and auditory processing ability.

3.4.4 Summary - Language Aptitude

Whatever mix of these possible variables, or of other unresearched areas, are included in the term *language aptitude*, some big claims have been made for this, as yet, ill-defined ability. In a study of 42 near-native speakers of Swedish, Abrahamsson and Hyltenstam (2008) assessed participants for various categories of aptitude (phonetic memory, lexical-morphological analytical skills, grammatical inferencing skills, aural memory for unfamiliar sounds and the ability to form sound-symbol associations) and proposed that “language aptitude is a necessary condition for eventually passing for a native speaker in normal everyday communication” (p.503). While researchers remain divided on whether there is a global concept of ‘language aptitude’, it is possible that certain learners have a facility with certain aspects of language acquisition. However, in the absence of clear evidence to the contrary, this research follows Moyer (2015a) in assuming that “improvement is arguably within anyone’s reach, and ‘talent’ may be less influential than one’s overall orientation” (p.10). In other words, as with so many skills, attainment appears to be a result not so much of aptitude, but of motivation and hard work.

3.5 GENDER

It has sometimes been claimed that females have greater aptitude for language learning than males. Van der Silk et al. (2015) carried out a large-scale survey of test results of 27,119 adult learners of L2 Dutch from 88 countries of origin with 49 different mother tongues and reported that, “we found a consistent gender effect for speaking and writing proficiency: Female learners outperformed male learners, independent of country of origin and mother tongue” (p.15). The gender effect for speaking seemed largely independent of other variables such as AoA, LoE, education and L1, but the authors noted that the data was not detailed enough to identify which components of speech contribute to this finding and whether pronunciation played a part.

Research into this area would appear to be sparse. Moyer (2016) comments that “evidence for gender effects in phonology is much like that for L2 more generally; very few studies *set out* to compare gender differences in rate, route, or ‘end-state’ of acquisition” (p.17), with findings often noted as a bi-product of other research. Piske et al. (2001) point out that, “most studies have not identified gender as a significant predictor of degree of L2 foreign accent” (see, for example, Snow & Hoefnagel-Höhle, 1977; Flege & Fletcher, 1992; Kassaian, 2011). Moyer (ibid.) notes that there is also a lack of any conclusive evidence from neurolinguistic research and that “based on what is known thus far, it is difficult to draw concrete conclusions” (p.16).

However, there is a considerable body of research which *has* found gender effects in phonological acquisition. Diaz-Campos (2004) found that, in a Spanish read-aloud task conducted with L1 English learners of Spanish, females outperformed males in terms of ‘nativelike pronunciation’. The author concluded that, “female students are more careful in producing the target segments, which can be interpreted as a conservative tendency” (p.268), the implication being that women are more concerned with conforming to standard norms of pronunciation, while this is less important to men. A similar conclusion is arrived at by Spezzini (2004), who noted gender differences in the English pronunciation of Spanish-speaking 12th graders at the American School of Asuncion, Paraguay. She concluded that “the variability in L2 learning between girls, approaching a more standard norm, and boys, following a non-native form, is best explained via both covert prestige ... and the Accommodation Theory” (p.426). Accommodation Theory (Giles et al, 1973; Giles & Powesland, 1997) looks at how speakers adjust their speech behaviours in accordance with their desire to identify with or distance themselves from interlocutor(s). Spezzini is suggesting that, in the school context described, non-native forms may be seen as having a higher prestige amongst male students and the boys therefore ‘accommodate’ towards those forms, while women ‘accommodate’ to more standard norms.

Hansen Edwards (2008) noted that perceived gender differences in pronunciation attainment might be a social construct in that gender may influence the quality and quantity of L2 to which the learner is exposed. In a longitudinal study of a Vietnamese couple who immigrated to the US, she found that the husband, while reporting fewer hours per day speaking English, spoke the L2 in informal contexts with English native-speaking friends, while the wife was exposed to more limited, formulaic exchanges in her job as a manicurist. Consequently, while the wife continued to have a marked NNS accent, the husband’s pronunciation improved over time. While this example gave the socially-influenced pronunciation advantage to the man, other research suggests that it is women who benefit more from social gender roles. Polat and Mahalingappa (2010), in a study of 121 Kurdish youths learning Turkish, found the girls were judged to be more native-like in spontaneous speech samples, which the authors related to girls having more widespread social connections, whereas the boys tended to socialise with peers of the same ethno-linguistic background. This led the authors to conclude that, for girls, “acculturation and integration into the L2 community is one of the integral contributors to their L2 accent” (p.31).

Social attitudes could also be at play here. Moyer (2010), in a survey of 42 learners of English in the US, found pronunciation-related gender effects which favoured women in both read-aloud and spontaneous speech production tasks. While there seemed to be no correlation with L2 use, desire to sound native, length of instruction or attitude to the language, the author noted that women were more likely to believe that accent improvement was possible and that they had a positive attitude to American

culture. Thus, it would therefore appear that the possible phonological advantage conferred on women may have more to do with socially-learned attitudes and behaviours than it does with biological differences. As Polat and Mahalingappa (ibid.) point out, this is an inevitable conclusion since “linking L2 attainment to gender regarding its biological characterization poses spurious conclusions because gender is socially constructed, and, therefore it constitutes social, cultural, and situational factors that shape its categories as well as roles” (p.29).

In conclusion, due to the mixed results obtained from research into the gender/pronunciation attainment link, along with the problematic nature of separating the biological from the socio-cultural aspects of what we mean by *gender*, it is very difficult to come to any meaningful conclusions as to its role in phonological attainment.

3.6 PERSONALITY

A further ID variable which has been linked to SLA in general and pronunciation acquisition in particular is the area of personality, the notion that certain personality traits could favour or hinder phonological acquisition. Much SLA research on personality has built on work into the Big Five taxonomy of personality (see Chen et al., 2021), namely: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism. While much of the research has focused on SLA in general, this section reviews studies which have focused on the role of personality in pronunciation acquisition, which have largely focused on two areas: extraversion and agreeableness (Moyer, 2015).

Extraversion has been described as a willingness to use the target language regularly (Moyer, 2015) and engage in social interactions which, it is proposed, maximise contact with the target language (Hu & Reiterer, 2009). Zárate-Sánchez (2017) tested 51 L1 English learners of Spanish using a picture elicitation task to rate pronunciation ability, while extraversion was assessed via interview and a questionnaire. He found that “personality explained a considerable portion of the variance, and that extraversion and neuroticism were significant predictors of foreign accent” (p.227). Neuroticism seems to equate closely with the concept of ‘language anxiety’ discussed later in this section. Similar claims for the role of extraversion in pronunciation aptitude are cited elsewhere in the literature. (See, for example, Hu & Reiterer, ibid; Ioup et al., ibid.). Furthermore, Hinton (2012) in his study of mimicry described in Section 2.4.2 above, notes that mimicry scores appeared to be influenced by extraversion or the lack of it - introversion. He states, “of the two candidates with the lowest scores for mimicry, one was somewhat shy in responding during the test” (p.109-110). While Hinton does not elaborate on this issue, the implication is that the candidate’s shyness was hindering performance.

The second personality trait which has often been mentioned in relation to pronunciation is *agreeableness*, the ability to understand how someone else is feeling (Gholamrezapour & Rabori, 2019) and to be helpful and sympathetic to others (Chen et al., 2022), often referred to as *empathy*. This relates to the idea of Accommodation Theory, introduced in Section 2.5 on ‘Gender’ above, which describes how people adapt (or accommodate) their speech and behaviour depending on their interlocutor. Relating this to L2 pronunciation, Guiora et al. (1972) proposed the notion of the ‘ego boundary’ and suggested that the more willing speakers are to change their language ego in verbal exchanges, the more they will be able to imitate NS pronunciation. In their conclusion, they state that their research “lends strong experimental support to the view that permeability in the language ego (specifically in pronunciation) will be a reflection of the permeability of ego boundaries in general and of empathic capacity in particular” (p.428), but the research has been heavily criticised due to the use of alcohol as a way to increase the ‘permeability’ of the ego boundary. However, subsequent research has supported the proposed link between empathy and phonological attainment. For example, Gholamrezapour and Rabori, (2019) studied pronunciation in 69 female L1 Iranian learners of English. Pronunciation was evaluated on a word read-aloud task, while empathy was assessed using Jolliffe and Farrington’s (2006) Basic Empathy Scale – a self-report questionnaire, and the authors found that “the high-level group (i.e. more empathic learners) were more successful in acquiring authentic pronunciation than those in the low-level group” (p.107).

While extraversion and empathy have garnered the ‘lion’s share’ (Moyer, 2021, p. 64) of attention in SLA research, other personality factors have also been implicated in pronunciation attainment.

Hu and Reiterer (2009) report on a study conducted with 62 German students of English, whose personality was measured with the NEO Five Factor Inventory (NEO-FFI, Costa & McCrae, 1992) and compared to ‘pronunciation talent’ in English as rated by 1 English NS. They reported a moderate correlation between talent and agreeableness, but also talent and low *conscientiousness*, a Big Five trait associated with reliability, good organization and self-discipline (Zárate-Sánchez, 2017), leading them to conclude that pronunciation talent was associated with being “more agreeable and less conscientious” (p. 120). Baran-Łucarz (2023) compared the personality traits of 58 L1 German learners of English as assessed by the Polish version of Goldberg’s battery (1999), with their degree of accentedness in English as assessed by 2 NS judges. She found a weak positive link between accentedness and the Big Five trait of *openness to experience*, generally defined in terms of intellectual curiosity and an openness to new ideas and experiences. Lewandowski and Jilka (2019) report similar positive correlation between openness and phonetic convergence, a willingness to accommodate to the interlocutor, which they found was a feature of phonetically talented individuals in their study. Phonetic convergence was also positively linked to *neuroticism*, defined as differences in the way we tolerate stress and its impact

on performance. The authors suggested that while neuroticism alone might not bring positive pronunciation outcomes, in conjunction with other personality traits such as openness and extraversion, it might promote pronunciation attainment.

While the Big Five relate to general personality traits, another language specific personality trait has been widely studied, that of *language anxiety*, (Clément et al., 1994; Krashen, 1982; Scovel, 1978). As it relates to pronunciation, language anxiety is defined by Baran-Lucarz (2016) in terms of fear of negative evaluation, negative self-assessment of pronunciation, poor pronunciation self-image and negative learner beliefs regarding pronunciation. This link to pronunciation attainment would appear to be supported by Hinton (2014) who tested L1 Polish speakers on their mimicry ability in Italian, Dutch and Chinese, which was then judged by native speaking raters for each language. He reported “both pronunciation and mimicry are negatively affected by high levels of anxiety” (p.217), but it is worth noting that anxiety was assessed by self-report questionnaires, which might be seen as subject to questions of validity. Indeed, Hu and Reiterer (2009) highlight this issue when they comment that “the definition of anxiety has been rather inconsistent in the body of SLA research” (p.107).

So, in the field of pronunciation attainment, various personality traits have been said to either help or hinder achievement, but it is important to note that this is an area of research fraught with difficulty. Hu and Reiterer (ibid.) comment that it is “not easy to carry out studies in the research field of personality and SLA ... partly because it is hard to find an appropriate level of analysis for both the personality and the language variables” (p.98). They go on to talk about the inevitable vagueness of defining and testing for a particular personality trait and determining to what extent they relate to each learner (see also Dörnyei, 2005). How does empathy relate to the idea of the ‘ego boundary’, and how does that, in turn, relate to the concept of ‘openness’? Personality is not a characteristic which can be neatly divided into discrete portions. Furthermore, it is difficult to tease apart the relative importance of these overlapping features of personality. If a learner is empathetic, but introverted and anxious, how do these apparently competing features manifest themselves in attainment (see Dörnyei, ibid.)?

It would appear that, at present, while it might be possible to make vague statements about the importance of personality in pronunciation aptitude, it is very difficult to separate the various facets of personality involved and to understand which combinations of these facets are promoting or hindering attainment.

3.7 ATTITUDE AND IDENTITY

Pennington (2019a) writes that “although it has long been assumed that changing the physiological basis of pronunciation – in brain structure and articulatory mechanics – is the most challenging as compared to the psychological and sociocultural bases of pronunciation, this may be an incorrect view” (p.375). It could be that the well-documented problems experienced by late L2 learners in pronunciation attainment lie, at least in part, beyond the level of segmental and suprasegmental competence and are, in fact, partly governed by an individual’s attitude to the L2 and degree of identification with the L2 speech community which, in turn, influence motivation.

Dörnyei (2009) has proposed a theoretical framework which he calls the “L2 Motivational Self System” to describe the different layers of motivation which influence L2 learner attainment. He talks about the often competing factors of the ‘ideal L2 self’ and the ‘ought-to L2 self’, the ideal L2 self being the learner’s vision of what they wish to achieve, while the ought-to L2 self describes external factors such as cultural and educational expectations, and professional demands, which might colour the ideal L2 self in some way. To this model, Dörnyei adds a third element, that of the L2 learning experience, which might also affect the previous two dimensions. This in agreement with other researchers (Moyer, 2015; Trofimovic et al., 2015) who situate motivation as a socially situated phenomenon, which “bridges personality and cognitive style” (Moyer, 2015, p.69).

In the following sections, these three elements of the L2 Motivational Self System: the ideal L2 self, the ought-to L2 self and the L2 learning experience will be discussed under the headings ‘individual motivation’, ‘group affiliation’ and ‘L2 learning experience’ respectively.

3.7.1 Individual Motivation

Dörnyei (ibid.) talks of the ‘ideal L2 self’ as “the person we would like to become” (p.29) in our language learning journey in terms of language skills, proficiency and identity. In this model, if the ideal L2 self is a proficient speaker of the L2, this provides strong motivation to reduce the gap between our current self and the ideal. Evidence for this notion comes from interviews with exceptional learners, who frequently mention a genuine interest in the target language and a strong desire to improve as factors in their learning (Abrahamsson & Hyltenstam, 2008; Ioup et al.1994; Moyer, 2014a; Piller, 2002;). Abrahamsson and Hyltenstam use the term ‘devotion to language’, which they describe as “a strong and often highly academic interest in issues related to language and language learning” (p.500). Saito et al. (2018), in a study involving 108 Japanese, high school, EFL students

found a correlation between their motivation, described as a “clear vision of ideal future selves” (p.1) and improvement in the comprehensibility rating of their L2 speech as evaluated by 5 NS English raters. Similarly, Nagle (2018) used a questionnaire based on Dörnyei’s L2 Motivational Self System to evaluate how motivation was linked to comprehensibility and accentedness of 26 L1 English learners of Spanish as rated by 18 NS Spanish listeners. No link was established between motivation variables and comprehensibility, but effort, defined as a subcomponent of motivation, was significantly related to accentedness. Moreover, motivation is not a stable trait, it might wax and wane over time and Moyer (2004), in a study of L2 German users in Berlin, found that their consistency of motivation over time correlated positively with pronunciation attainment.

A lack of motivation might work in the opposite direction to hinder the achievement of nativeness. This could be linked to the learner’s previously held views on limits to the end state of attainment, which lead the learner to believe that accent modification is just not possible if learning starts after a certain age (Flege & Mackay, 2011; Moyer, 2015; Muñoz & Singleton, *ibid.*; Piller, *ibid.*).

However, it is very difficult to separate individual motivation from social context. Saito et al. (2018) talk of the ideal L2 self as including both *integrative motivation*, the desire to integrate into the L2 speaking community and culture, and *instrumental motivation*, a desire to learn for the tangible rewards of speaking the L2 well, such as better job opportunities or financial rewards, meaning that the social context of the learner dictates, to a large extent, the ideal L2 self. Pennington (1997, p.80) talks of this in terms of the *individual self* and the *collective self*. The individual self refers to personal goals and motivations of the learner, as discussed above but, as Pennington points out, this is very much shaped by the social environment in which the learner is operating: the collective self.

3.7.2 Group Affiliation

Dörnyei (*ibid.*) describes the ‘ought-to L2 self’ as possessing the qualities the learner believes they should possess in order to meet listener expectations. Moyer (1999), citing research into advanced L2 German speakers whose L1 was English, comments that the learner whose accent was most highly rated by NS judges professed “a strong desire to acculturate and to sound German” (p.98). Other studies have identified a similar desire to assimilate into the L2 culture as a factor influencing degree of L2 foreign accentedness (Gatbonton et al. 2011; LeVelle & Levis, 2014; Lybeck, 2002; Moyer, 2004). As described above, this might involve trying to adhere as closely as possible to NS practices, but it might equally involve distancing oneself from those practices, in order to signal one’s identity as an NNS. Derwing and Munro (2005) note that an NNS accent “signals to an NS interlocutor than an L2

learner is non-native and may therefore require modified input” (p.385) and Piller (ibid.) reports on a Danish speaker of German who self-reported passing for a German but “prefers to make it clear very early in an interaction that she is not native” (p.195) in order to avoid socio-cultural misunderstandings. In a more extreme case, Moyer (2015b) reports on a study of L2 German speakers living in Germany in which two of the participants “expressed an actual aversion to ‘sounding German’ or being too closely affiliated” (p.4). Research has suggested that accent is the strongest linguistic marker of a speaker’s cultural identification (Guiora, Brannon & Dull, 1972; Labov, 1972; Scovel, 1978). It would seem, therefore, that accent goals are closely tied to learner attitudes towards the L2 and the individual and social identity which the learner wishes to present to interlocutors.

Furthermore, explicit in Dörnyei’s (ibid.) description of the ought-to L2 self is a desire to avoid possible negative outcomes of not meeting listener expectations. That there are social costs involved to learners’ accents was discussed in Section 1.5.3 above in relation to possible NS prejudice with regard to NNS accents, but it is suggested the NNSs also make judgements about NNS accents. A study by Gatbonton et al. (2005) which collected L2 English speech samples from L1 French and L1 Chinese in Quebec and asked peers to judge their *ethnic group affiliation* (EGA) – the strength of their sense of belonging to their L1 ethnic group. The authors concluded that, “learners treat their peers’ L2 accent as an indicator of the peers’ degree of ethnic affiliation” (p.489). Pronunciation, therefore, is not just a reflection of the way the speaker chooses to present themselves to the world, but, like appearance and behaviour, is also a factor in how they are judged by others, including their peers. It is easy to understand that, whilst for some there is no conflict between the ‘ideal L2 self’ and the ‘ought-to L2 self’, for others, a desire to signal one’s individual or group identity through accent might lead to conflicting goals. Gatbonton et al. (ibid.) describe this choice in terms of a compromise “between the reward of being efficient and the cost of not marking identity” (p.489).

Some researchers (Derwing & Munro, 2009; Piller, 2002) have even suggested that having an NS accent might not be an ‘all-or-nothing’ ability. Piller (ibid) cites examples of late learners who claim to be mistaken for an NS in certain situations, an ability which Piller labels as ‘*passing*’. Passing is seen by her as a temporary skill, which the learner might use in certain situations. In this view, NS “is no longer an identity category, and rather than being something that someone is, it becomes something that someone does” (p.201).

3.7.3 L2 Learning Experience

The third aspect of Dörnyei's (ibid.) L2 Motivational Self System, the 'L2 learning experience', is defined by Dörnyei (ibid.) as the immediate learning environment and experience of the learner, including influences such as teacher, curriculum and peers, along with the learner's attitude to the learning experience.

It has already been discussed in Section 2.7.1 how erroneous learner beliefs about L2 pronunciation attainment might cloud their progress and in Section 2.7.2 that learner attitudes to the L2 itself and to the L2 community might encourage or deter. It is also possible that in an EFL context, where learners and/or the teacher share an L1, learners "converge toward second language (L2) pronunciation that is heavily influenced by the L1" (Levis, 2005, p.371). This may be the result of pressure to conform with L1-influenced norms, where learners might reject pronunciation 'standards' in favour of L1-accented pronunciation in the L2, in order to fit in with the majority (see Jenkins, 2007). Indeed, Jenkins (2000, p.80) describes an English information exchange task between an L1 Japanese speaker and an L1 European Portuguese speaker and noted how both subjects tended to drop final *\n* and nasalise the previous vowel sound in spontaneous speech, but in an information exchange task where, Jenkins notes, intelligibility was paramount, instances of this non-standard pronunciation feature were minimal. However, in the same study, where participants shared an L1, the non-standard feature increased from the spontaneous interaction to the information exchange task. It would appear that the learners in this study perceived a need to reduce the L1 influence in English interactions with speakers from a different L1 background, but chose to emphasise it when speaking to speakers of the same L1. This seems to be another example of Accommodation Theory, introduced in section 2.5 above, suggesting that speakers 'accommodate' their language to the interlocutor.

That this convergence might be working below the level of consciousness is suggested in a study by Lefkowitz and Hedgcock (2002) which talks about the "psycho-social barriers to learners' success" (p.224) that learners face in the classroom. In qualitative data obtained from interviews and questionnaires of 282 L1 English learners of Spanish and French, the authors concluded that, "social and affective variables such as desire for peer acceptance might create a substrate of the classroom FL learner population" (p.240), which they describe in terms of a 'covert', L1-influenced pronunciation goal. Furthermore, they comment on, "an inadvertent and often unacknowledged desire to underperform phonetically in the presence of classmates and teachers" (p.223), despite respondents professing standard pronunciation as their L2 goal. According to the authors, therefore, there is evidence that the learners themselves are not aware of their 'covert' pronunciation and show a general tendency to over-evaluate their oral performance (as predicted by the *Dunning Kruger effect*, Section

2.4.1.4 above). The authors also note that this ‘covert’ pronunciation effect is stronger among secondary school students than in mature students, suggesting that peer-influenced convergence is stronger during adolescence than adulthood.

3.7.4 Summary -Attitude and Identity

Birdsong (2006), in accounting for individual differences in native-like attainment in late learners, comments that, “it is pointless to invoke neurobiological capacities (or deficiencies) in the context of an individual who has no interest in passing for a native” (p.36). In other words, it is important not to assume that all learners want to sound like an NS. Gatbonton et al. (2011) also highlight this point when they note, “L2 users’ failure to demonstrate high levels of L2 oral proficiency do not necessarily imply a lack of ability to learn” (p.201). There might also be sociocultural factors which, consciously or sub-consciously, are influencing the learner to speak in a certain way. Pennington (2019a) talks of the need to move away from the notion of the language learner “at the mercy of L1 transfer” (p.374) to a view of the learner who has, at least to a certain degree, agency over their L2 pronunciation outcomes. She sums up this important role played by attitude and identity in pronunciation outcomes as follows:

“...changing pronunciation means altering psychological, social, and cultural aspects of the self. Pronunciation change is therefore not only a matter of ability; it is also a matter of openness and willingness to change based on valuing such change” (p.375).

3.8 CONCLUSION - FACTORS INFLUENCING ULTIMATE ATTAINMENT IN L2 PRONUNCIATION

While pronunciation acquisition has sometimes been seen as restricted by age-related factors, it is becoming increasingly accepted that the reality is more complex. Moyer (2014a) talks of “a constellation of influences on accent” (p.19) and it has been discussed how these include age-related limitations, but also other experiential factors, namely L1 interference and quantity/quality of exposure, along with the ID variables of language aptitude, gender, personality and identity. It is also important to be aware that these factors do not work in isolation, but work together to influence phonological output and their effects may show considerable individual variation. What is more, UA is probably not

a fixed category. A learner might be judged ‘native’ in some contexts but not in others and might even choose to sound more or less native depending on the situation.

Importantly for this research, while some of these factors are immutable (age of onset, gender, personality) others have been shown to be modifiable (language exposure, negative L1 transfer, phonological self-awareness, motivation, identity and, possibly, cognitive control) and nowhere have these factors been shown to be insurmountable given the appropriate strategies. The following chapter looks the effectiveness of various types of intervention in modifying pronunciation.

3. INFLUENCE OF INSTRUCTION ON L2 PRONUNCIATION ATTAINMENT

"There is clearly no single path to becoming skilled in any cognitive task, and learning to perceive and produce the sounds (phonology) of a second language (L2) like native speakers of that language is no exception" (Trofimovich & Gatbonton, 2006, p.519)

In the previous chapter it was noted that some exceptional late learners are judged to achieve native-like pronunciation in an L2 purely from exposure to the language. However, for the majority, this is not the case and it has long been debated whether L2 pronunciation instruction (PI) can go some way to overcoming this limitation. De Graaf and Housen (2009) define L2 instruction as "any deliberate attempt to promote language learning by manipulating the mechanisms of learning and/or the conditions under which these operate" (p.726) and a similar definition can be offered for PI. Pronunciation instruction involves a deliberate attempt to promote pronunciation modification in an educational setting. Thus PI, as discussed by Thomson (2014), is distinct from medical models, in which a speech/language pathologist aims to treat patients with speech disorders and also distinct from so-called accent reduction programmes, which are not necessarily taught by trained teachers or experts in the field and generally follow a business, that is a profit-making, model. The chapter starts with an overview of theoretical models of PI, followed by sections on different types of intervention, namely perception training, production training and global strategies to modify accent, that is, those which focus on overall performance rather than specific structural features. It finishes with a review of current attitudes to the teaching of pronunciation.

3.1 THEORETICAL MODELS OF PRONUNCIATION INSTRUCTION

In Chapter 2 above, L2 pronunciation acquisition was shown to be a complex process with many influences on eventual outcomes. As Grant (2014) points out, pronunciation instruction needs to "address the many levels on which pronunciation change occurs" (p.28). Early learners seem to learn their L1 through *implicit learning* processes (Bley Vroman, 1989; DeKeyser, 2000; Hulstijn, 2015, Dollman et al., 2020), that is, from unintentional acquisition without conscious reflection (Dollman et al., *ibid.*). However, for late learners acquiring an L2, varying outcomes suggest a breakdown of this learning process at some point. This could be at the level of awareness, perception, articulation, or possibly, be influenced by psycho-social pressures, related to personality/identity issues. Thus, for example, if a learner is struggling with the contrast between the voiced liquids /r/-/l/ in English, this could be for several reasons. The learner might be unaware of the existence of the /r/-/l/ contrast

(knowledge), unable to perceive the distinction in their own or others' productions (perception) or they may be physically unable to voice the phonemes in a nativelike way (articulation), or a combination of these factors. Alternatively, or additionally, conscious or unconscious psycho-social processes could be interfering with the process, such as language anxiety or phonetic convergence with L2 peers.

Schmidt (1990, 2001) proposes that, to overcome this late-learner disadvantage, adults have to consciously pay attention to language features which are not acquired implicitly in other words, they become more dependent on *explicit learning*, deliberate learning involving an awareness of what is being learned in order for it to be acquired. He formulated the *Noticing Hypothesis* which proposes a distinction, first suggested by Corder (1969), between *input* – the language which the learner is exposed to – and *intake*, what the learner actually attends to or 'notices', and claimed that, for late learners, "input does not become intake for language learning unless it is noticed ..." (Schmidt, 2012, p.721). While, for early learners, 'input' appears to become 'intake' when they are naturally exposed to the language triggering the creation of implicit knowledge, for late learners, this may occur only in some instances. When this is not the case, learners may need help 'noticing' input, for it to become 'intake' or, as Schmidt (2010) puts it, "some kinds of learning require more focused attention than others" (p.728).

For 'noticing' to be fully effective, it must apply not just to the input to which the learner is exposed, but also to learner output, so that both learner perception and learner production are being addressed. Schmidt (ibid.) talks of 'noticing the gap' – the concept that "to overcome errors, learners must make conscious comparisons between their own output and target language input" (p.724). Again, this may be problematic for the learner. In section 2.4.1.4, it was suggested that late learners may struggle to accurately evaluate their own performance in pronunciation and therefore are not in a position to work towards self-correcting. This being the case, pronunciation would appear to be an area of language which, for some learners, would benefit from instruction in order to help them 'notice the gap' between target input and their own output.

Furthermore, while pronunciation has a knowledge component, a knowledge of phonetics and phonotactic rules, it is largely a physical skill (Archibald, 2018; Fraser, 2006; Moyer, 2013; Richards & Colos, 2019). Skill Acquisition Theory posits three stages of development (DeKeyser & Criado, 2013; DeKeyser, 2015) moving from *declarative knowledge* or 'knowledge about' a skill, which might be acquired by exposure or explicit teaching, through to *procedural knowledge*, or 'knowledge how' to perform the skill, acquired through practice. However, as DeKeyser (2015) points out, "once procedural knowledge has been acquired, there is still a long way to go before the relevant behavior can be consistently displayed with complete fluency or spontaneity" (p.95) in other words, before

automatization of the skill. For procedural knowledge to become automatized, according to this model, requires “a large amount of practice” (p.95). This has important implications for pronunciation acquisition and modification since, as Richards and Colos (ibid.) note, “pronunciation, more than other language skills is necessarily nearly always produced automatically” (p.416). In other words, pronunciation is an online skill, in which the speaker has little or no time or processing capacity to consciously devote to pronunciation. Situating pronunciation as a skill which requires automatization, therefore, has consequences for the classroom. Firstly, as Fraser (ibid.) comments, it cannot just involve the passive reception of knowledge, but it requires active engagement on the part of the learner. Secondly, it requires considerable practise (Richards & Colos, ibid), especially for learners beyond the beginner level, who might be working to “*unlearn* previously acquired (i.e. now fossilized [see Section 2.3]) mispronunciations” (p.416). Richards and Colos (ibid.) suggest that working 10-15 minutes per day on building new segmental pronunciation habits is likely to ‘soon’ (p.418) lead to modification. Furthermore, Fraser (ibid.) argues that practising in class is not nearly as important as practising outside class and the job of the teacher is to motivate students to continue their practice after they have left the classroom.

The following sections look at the relative effectiveness of different forms of pronunciation intervention in the modification of accent: perceptual training, production training and lastly, the training of *global strategies*, more holistic, strategy-led training to promote pronunciation modification.

3.2 PERCEPTUAL TRAINING

3.2.1 Perception Training Research

There has been considerable interest in the role of perception in the field of L2 pronunciation acquisition and particularly the notion introduced by Strange (1995) that, while speech production is accented, perception may also be ‘accented’. As discussed in Section 2.2 above, Flege’s (1995) Speech Learning Model (SLM) proposes that perception precedes production and that, if an L2 phoneme is perceived as similar to a learner’s existing L1 sound, the learner might fail to recognise it as a new speech sound and instead integrate the new L2 sound within an existing L1 category. This, in turn, may lead to ‘accented’ production. Part of the role of PI in this paradigm is to help the learner to ‘hear’ the difference between the two sounds and, as a consequence, perceptual training has tended to focus on minimal pair work.

Considerable success has been achieved in this area, particularly through the use of computer assisted pronunciation training (CAPT) in general and, more specifically, high variability phonetic training (HVPT). HVPT exposes the learner to a wide range of examples of a particular sound or contrast produced by multiple talkers in different phonetic contexts, generally using both identification and discrimination tasks with feedback for the participants. An example is a study by Rato (2014), which used HVPT in an attempt to improve the perception of three American English vowel contrasts (/i-ɪ/, /ɛ-æ/ and /u-ʊ/) by 34 L1 European Portuguese speakers who were intermediate-level (B1) EFL learners. While /i/, /ɛ/ and /u/ are present in the L1, the other three vowels are not and therefore, as predicted by the SLM, might be integrated into the vowel space of the L1 categories. The results suggested that “perceptual training significantly improved the Portuguese learners’ performance in the identification of the three target vowels” (p.542). Similarly, Bradlow et al. (1997) used HVPT to train Japanese learners of English in the perceptual identification of the voiced liquid /r-/l/ contrast. The learners were between 19 and 22 years of age and had studied English since +/- age 12. Not only did they find an improvement in perceptual learning, but this improvement also transferred to the production of minimal pairs containing these phonemes.

However, the Bradlow et al. (ibid.) study seems to be somewhat of an outlier when it comes to showing a strong link between perception training and production. Sakai & Moorman (2018) in a meta-analysis of 25 years of research into the role of perception training in production comment that, “overall, the studies show that perception training is largely successful ... However, perception training studies have not been able to reliably show improved production” (p.5). For example, Qian et al. (2018) used the HVPT paradigm to produce customised perception training exercises for 32 adult L1 Russian learners of English, but found that perceptual gains in target sounds did not transfer to untrained words. Similarly, Dufour et al. (2010) found that the successful training of Southern French listeners on the standard French /e-/ɛ/ contrast, which is not contrastive in their dialect, did not generalise to regular use of this contrast in spoken word recognition tasks.

Indeed, it is possible that the importance of the perception/production link has been over-emphasised, as suggested by a recent study by Kirby & Giang (2021). The authors tested speakers of Khmer, a non-tonal language, in their ability to discriminate and produce Vietnamese tones. They found that participants’ tonal perception could be quite poor, even when their production of the tones was considered native-like and the reverse, that listeners could be relatively good at discriminating the tones, even when their productions were not native-like. While recognising that their findings might have been a result of task type, the authors felt able to conclude that the findings suggest that “L2 production can be independent of perception” (p.268). Nagle et al. (2022) suggest that the apparently

contradictory results of research into the relationship between perception and production may be the result of differences in research design and call for more consistency in research methodology.

Even if perceptual training is shown to transfer to production, there may be other obstacles to integrating it into the language classroom. Thomson (2018b), while describing HVPT as “a very effective pronunciation training tool” (p.208) also points out that, at present, HVPT applications seem to focus on a limited number of phoneme contrasts. For HVPT training to be more widely used in SLA, it would have to encompass the entire phonology of a given language, which “soon adds up to thousands of tokens, which must be recorded and screened for accuracy” (p.226). It is unlikely that such an extensive project could be provided without financial costs to the learners. Thomson also offers what is possibly a more serious criticism of this sort of repetitive training when he notes that current HVPT training paradigms tend to be ‘boring’ (p.227) for participants and therefore unmotivating. He suggests that a possible solution to this shortcoming is the integration of HVPT within a computer game, such as the small-scale study by Lim and Holt (2011), but again this may be costly and thus inaccessible in the majority of teaching contexts.

Therefore, while HVPT has proven effective in achieving perceptual gains, it is not clear whether these gains transfer to production and, even if they did, how this could be made simple, engaging and accessible for use in the classroom.

It has also been questioned whether a focus on segmental training is useful. Hardison (2003) questions “the value of segmental-level perceptual training using words spoken in isolation because such stimuli are devoid of the sandhi phenomena typical of the connected speech generally encountered” (p.517-8). There appear to have been few studies which address perception above the level of the segment. Rato and Oliveira (2022), in a meta-analysis of 27 perceptual training studies, talk of “the underrepresentation of suprasegmentals” in their dataset with “the vast majority of training studies focusing on segmental categories” (p.291). (See also Thomson, 2022.) Rato and Oliveira (ibid.) also make the important point that their meta-analysis might reflect a publication bias, in the sense that research that fails to identify a positive learning outcome is often rejected from publication, the so-called ‘*grey literature*’ (p.294). The authors highlight the need for the inclusion of all relevant studies, not just those which show significant results, in order to make informed judgements on the effectiveness of speech perception training.

There have also been those who argue that perception training needs to include, not just the perception of external stimuli, but should include perception of the self. In other words, learners need phonological self-awareness of how others are perceiving their accent and this is the topic of the following section.

3.2.2 Phonological Self-Awareness Raising Activities

In section 2.4.1.4, it was suggested that learners' evaluation of their own accent is often subject to the Dunning-Kruger effect, that is, they show a tendency to inflate their self-assessment of performance in oral tasks (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al., 2016). Couper (2015) proposes that, as the point of departure for any PI, "we have to understand the difference between what we actually say (the physical sounds) and what we think we say (the phonology, or the way we categorize sounds)" (p.421). Similarly, McGregor and Reed (2018) describe an awareness of pronunciation needs as among the primary steps to address in order to provoke pronunciation modification. There is now growing evidence that such self-appraisal can be taught and various strategies have shown success in the literature.

In a study by Lappin-Fortin and Rye (2014), learners of L2 French were asked to assess their own oral production several times over the course of a 12-week programme into certain segmental and prosodic aspects of the French sound system and then compared their self-assessments with expert assessments at various stages of the intervention. The researchers reported that the feedback prompted learners to produce more native-like pronunciation in their speech and "their awareness of their pronunciation difficulties had increased" (p.300). Couper (ibid.) advises that learners should "record themselves at the beginning of a course and use this as a diagnostic" (p.426) and there is evidence that this is effective in promoting pronunciation modification. MacDonald (2018) reports on a study in which he recorded students' spontaneous speech "on a topic that was relevant or important to them" (p.255) and then, with the students' approval, re-recorded the content himself to provide an NS model. The two were then compared, with students invited to consider which features of the NS model "might be something they could adopt to enhance their own speaking clarity" (p.257). While not reporting research evidence, MacDonald reported positive outcomes for the procedure. Similar success is reported by Couper (2003), who describes a PI intervention course, which included a task for learners to monitor their own pronunciation by recording themselves after listening to a model and then relistening and comparing. While there is no evidence that it was this specific task led to the improvements achieved, in a post-course feedback questionnaire, it was the task which received the highest score for 'how far the activities ... were considered helpful' (p.62).

There is also evidence that asking students to reflect on their progress through diary entries or questionnaires elicits positive results. For example, in a 13-week university-level PI course, Kennedy & Trofimovich (2010) compared learner language awareness diary entries by the learners with the quality of their pronunciation as assessed by NS listener ratings. They found a positive relationship between the number of qualitative comments, described as comments referring to "language learning as

a meaningful context in which learning occurs” (p.171) and their perceived pronunciation improvement, suggesting that their level of language awareness was helping them to improve. Meritan & Mroz (2019) found that learners of French as a Foreign Language improved their production of the standard French /y/-/u/ distinction more when self-reflections were included along with explicit instruction, compared with explicit instruction only. Meanwhile, Sardegna & McGregor (2013) asked learners to self-assess their pronunciation progress through questionnaires. The authors champion the choosing of pronunciation goals based on student needs, but as participants in the study were training to be International Teaching Assistants (ITAs) at the University of Texas, much of the material focused on GA as the target, with American cellist, Yo-Yo Ma used as the model speaker in one activity. While findings were hampered by having no control and no delayed post-test measures, results seemed to show that encouraging students to self-assess their own production was motivating for them and spurred them to propose ways of modifying features which they identified as needing modification (see also Schmidt & Frota, 1986).

While the self-assessment tasks above generally involved teacher monitoring and feedback, Miller & Ng (1994) found that involving other students in this assessment process was an effective activity in achieving pronunciation modification gains and, importantly, they reported that the reliability of the peer assessment was comparable to the teachers’ assessment. Tsunemoto et al. (2022) found a similar positive effect of peer assessment in a study involving learners of French in which learners were invited to evaluate their peers’ speaking performances. The authors concluded that “peer-assessment activities were generally effective at narrowing the gap between L2 French learners’ self-ratings of comprehensibility and the evaluations of comprehensibility by the external raters” (p.151). Involving peers in the assessment process, as well as showing a positive outcome, also means that learners have various sources of feedback on their own production and helps to create a learner-centred classroom, where students are involved in the learning process.

While more work needs to be done to better understand the perception/production relationship, to look at the role of perception in the training of suprasegmental features, such as rhythm, word and sentence stress, intonation, elision, contraction and assimilation, and to produce engaging teaching materials to facilitate learning, perception training alone would appear to be insufficient in addressing many pronunciation difficulties and a focus on production may be necessary. For example, Herd et al. (2013) point out that many English NSs have difficulty producing a Spanish alveolar trill /r/, despite the fact that they can hear it. The following section looks at research findings into the effectiveness of different types of production training in accent modification.

3.3 STRUCTURE-FOCUSED PRODUCTION INSTRUCTION

3.3.1 Focus on Form versus Focus on FormS

Form Focused Instruction (FFI) is the name given to any instruction which purposely focuses the learner's attention on a particular aspect of language. Long (2000) distinguishes between 3 types of intervention: Focus on Meaning (FonM), Focus on Forms (FonFS) and Focus on Form (FonF). FonM involves communicative activities without overt attention to specific forms of the language which, as noted in Section 3.1 above, may be insufficient to lead to 'noticing' in late learners in the case of pronunciation. FonFS describes intervention, based on a linguistic syllabus, where the teacher systematically and explicitly introduces and practises discrete areas of language predicted to be necessary for the learners. Ellis (2016) describes it as "a structure-based approach ... where linguistic forms are taught directly and explicitly" (p.405). FonF describes communicative, meaning-focused instruction, where the focus on language forms is secondary to the focus on meaning. Ellis et al. (2002) identify two forms of FonF: *planned FonF*, designed to provoke the use of a specific linguistic form but in the context of a meaning-focused task and *incidental FonF*, where the focus on form is incidental in response to learner needs, within a meaning-focused task. The former has been referred to as *proactive* attention to form, while the latter is *reactive* (Loewen, 2011). While planned FonF bears similarities to FonFS, it differs in that linguistic forms are only addressed once the learner has already been exposed to the use of the form via a meaning-focused activity. Dougherty & Williams (1998) state that "the fundamental assumption of focus-on-form instruction is that meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across" (p.4).

There are researchers who advocate an FonFS approach to PI. For example, DeKeyser (1998) proposed that "pronunciation is relatively immune to all but the most intensive forms-focused treatments" (p.43), but research has tended not to support this view. In a meta-analysis of 15 pronunciation intervention studies, Saito (2012) found, of the 6 studies using FonF instruction, all showed improvement at the controlled level and 2 tested for and showed improvement at a spontaneous level. However, the 8 studies using FonFS instruction showed improvement in controlled tasks, but none of the 3 of those studies which measured for transfer to spontaneous productions showed any improvement. The author concludes that "whereas FonFS tends to lead to improvement only at a controlled level (e.g. Elliott, 1997; Saito, 2011), FonF enables learners to achieve improvement both at a controlled and spontaneous levels (Derwing et al., 1998; Saito & Lyster, 2012)" (p.815).

Similar results have been reported elsewhere (Trofimovich & Gatbonton, *ibid.*; Isaacs, 2009; Tabandeh et al., 2018). So, evidence seems to support the idea that the largest changes are to be achieved using an FonF approach to instruction in which the focus is on meaning and any explicit focus on pronunciation is in response to learner needs as identified.

However, Ellis (2016) has questioned whether the distinction between FonFS and FonF really is such a clear-cut one or whether it can be better seen as a question of procedures, rather than being two completely different approaches. He notes that “explicit language teaching [FonFS] can also include activities designed to focus learners’ attention on form in communicative activities” (p.409), such as in the production phase of the traditional presentation-practice-production (PPP) teaching paradigm. This has led the author to describe the FonF/FonFS distinction as being ‘a continuum’ (p.410) from a meaning focus at one extreme to explicit attention to form at the other. This is a view supported by Doughty & Williams (*ibid.*) who state that the two “are *not* polar opposites” but that “FonF *entails* a focus on formal elements of language, whereas FonFS is *limited* to such a focus” (p.4). PI, to be effective, would appear to need both an attention to form *and* meaningful communicative practice, which Isaacs (*ibid.*) talks of in terms of “striking the right balance between form and meaning” (p.1). She refers to the work of Firth (1992), who proposed the ‘zoom principle’ where pronunciation work would constantly shift from communicative activities to a specific diagnostic pronunciation focus, back to a communicative focus, and so on.

It would appear, therefore, that in production training, research suggests that an attention to form within a communicative syllabus is the most effective intervention to achieve accent modification. The following section will address the question of what features of pronunciation should be the focus of instruction.

3.3.2 Segmental versus Suprasegmental Instruction

In terms of the content of pronunciation intervention, treatments have generally focused either on segmental or on suprasegmental aspects of production. However, Thomson and Derwing (2015a), in a narrative review of 75 L2 PI studies, comment that “suprasegmental research is under-represented” (p.4) and this, despite that fact that research seems to point to suprasegmental features as being at least as important as segmental in accent modification. A study by Derwing et al. (1998) compared three groups of learners: one, the ‘segmental’ group, received training in segmental accuracy, a second, the ‘global’ group, received prosodic training and the ‘control’ received no explicit pronunciation instruction. They found that both the treatment groups were judged to be ‘less accented’ in a sentence

read-aloud task, but only the ‘global’ group were judged to have improved in spontaneous speech. The authors proposed that, in spontaneous speech, the attention of the participants was taken up with lexicogrammatical concerns and proposed:

“speakers could not allocate enough resources to phonological concerns for there to be a noticeable transfer of segment-based skills. On the other hand, speakers who had had instruction emphasizing prosodic features such as rhythm, intonation, and stress could apparently transfer their learning to a spontaneous production.” (p.406).

Other research has supported this finding that suprasegmental training is more effective than segmental in accent modification (Missaglia, 1999; Fraser, 2001; Gordon et al., 2013; Gordon & Darcy, 2016; Gordon & Darcy, 2019), with different suprasegmental dimensions being highlighted (prosody, Missaglia, *ibid.*; word and sentence stress, Fraser, *ibid.*; primary stress, Hahn, 2004; epenthesis and absence, Couper, 2006; nuclear stress and final intonation, Muller & Levis, 2016). Similarly, Zhang and Yuan (2020) found that both segmental and suprasegmental training were effective in improving the comprehensibility ratings of 61 Chinese intermediate level ESL learners as judged by 6 NSs in a sentence-reading task. However, only the effects of suprasegmental training transferred to spontaneous speech in this study.

However, while these findings emphasise the importance of suprasegmental PI, nowhere does there appear to be an appetite for the complete abandonment of segmental training. Fraser (*ibid.*) comments that “there is little point in helping students with, say consonant distinction, if they have very poor control of word and sentence stress” (p.34), but she nevertheless includes consonant and vowel production in her list of pronunciation features to be taught once suprasegmental features have been addressed. Similarly, Gordon & Darcy (2016), while finding a group instructed in vowel sounds alone (segmental features) were less comprehensible than prior to instruction, warn against abandoning segmental instruction altogether. Instead, they argue that “an explicit pronunciation curricular component, integrated in oral communication classes — including both segmental and supra-segmental pronunciation features — can significantly improve comprehensibility, even in a short period of time” (p.84).

Indeed, rather than seeing segmental training as the poor relation, there is some evidence that it might play a significant role in listener judgements. Sereno et al. (2016) compared the role of segmental features versus intonation on accentedness, comprehensibility and intelligibility ratings by NSs and found that “segmental information contributes substantially more to the perception of foreign accentedness than intonation. Native speakers seem to reply mainly on segments when determining foreign accentedness” (p.303). Similar findings are reported by Saito et al. (2016).

Alternatively, there are those who say that the apparent segmental/suprasegmental distinction is a chimera since the two are indivisibly linked and that it is important to look at both segments and suprasegments together when understanding what happens within coherent stretches of speech or *thought groups* (Celce-Murcia et al., 2010), rather than looking at either in isolation. For example, Grant (2014) compares the utterances ‘*Leave me aLONE!*’ and ‘Don’t LEAVE me!’ and notes that, in the former, the word *leave* is not prominent and therefore likely to be assimilated to “*leamme alone*” (p.27), while in the latter, *leave* carries the nuclear stress and is therefore likely to be in citation form. Thus, the segmental feature, /v/ disappears in the first phrase due to its lack of prominence within the thought group, a suprasegmental feature and the two elements cannot be understood in isolation. This is supported by Lee et al. (2015), who, in a meta-analysis of 86 L2 PI studies, reported on “the larger effects observed when PI targeted both segmental and suprasegmental features (as opposed to either one independently)” (p.361). That the debate around the relative merits of segmental vs. suprasegmental instruction may be unhelpful is supported by Saito (2012), who, in the meta-analysis previously mentioned, commented that “teachers should carefully select target features for their own students ... instead of preoccupying themselves with the debate on the competing merit of segmental- vs. suprasegmental-focused instruction” (p. 815).

So far, it has been argued that production training within a communicative syllabus is more effective than a sole focus on structure and that both segmental and suprasegmental features need to be addressed, since it is questionable to talk of the two in isolation. The following section looks at what types of interventions have proven to be more effective.

3.3.3 Teaching Methodology

Thomson & Derwing (2015a), in a review of 75 L2 PI studies referred to above, report that 82% of the research they surveyed reported successful outcomes. This finding needs to be viewed with a certain degree of caution given the previously mentioned publication bias against *grey literature* showing no or negative effects, but the success of PI has been noted elsewhere in the literature (Saito, 2012; Lee et al., 2015), so appears to be a robust finding.

Unfortunately, studies often seem to be vague regarding the details of their intervention. In their meta-analysis, Lee et al. (ibid.) state that they were “often left wondering about the details of instructional materials and activities” and that “future studies would do well to include greater procedural detail in written reports (p.361). This was also a finding in the narrative review by Thomson & Derwing (ibid.), which noted that “the precise nature of instruction was often under-defined, making interpretation of

results and replication difficult” (p.5). It is therefore often difficult to understand the exact nature of the methodology used in successful interventions, but the following sections provide an overview of materials and activities which have been mentioned in successful research.

3.3.3.1 PPP Approach

Several of the studies which report more detail on their methodology refer to using materials from the seminal pronunciation textbook by Celce-Murcia et al., *Teaching Pronunciation: A course book and reference guide*, originally published in 1996, but later updated in 2010. For example, Gordon & Darcy (2016), in a study of the relative merits of segmental versus suprasegmental training, describe their treatment as including “decontextualized recognition and discrimination tasks, minimal-pair drills, reading of short passages and sentences out loud” (p.67), for which they give Celce-Murcia et al. (2010) as a reference. Other research which reports positive outcomes using this text in their methodology include Derwing et al. (1998), Alves and Magro (2011) and Gordon et al. (2013), while other studies which have shown improvement as a result of PI cite the work as a reference, such as Sardegna (2011), Couper (2011), Kissling (2013), Castelo (2017) and Carlet and Souza (2018). The materials in Celce-Murcia et al. (ibid.) take a broadly PPP approach. For segmentals, the PPP format starts with perceptual tasks such as identifying specific consonants within words (CD 1, Track 6) or more global tasks such as identifying consonant ‘problems’ in spontaneous NNS speech samples (CD1, Track 12). Production activities include sagittal section diagrams, plus details on place and manner of articulation. Phones are then often dealt with contrastively with perception and production minimal pair work activities leading to controlled and free practice tasks. For example, work on the /l/-/ɹ/ contrasts starts with listening discrimination sentences, such as ‘*the teacher collected/corrected the homework*’ (p.68), followed by a controlled practice roleplay with target sounds underlined:

Ron: *What’s the problem, officer?*

Police Officer: *Didn’t you see the traffic light?* (Celce-Murcia, 2010, p.68)

There are then suggestions for less controlled activities, such as being given a series of words containing the target sounds and being asked to create a story using eight of them (p.70).

Suprasegmental features follow a similar pattern. For example, the section on *sentence stress and rhythm* starts with a listening discrimination task using nursery rhymes (p.213), followed by controlled practice exercises such as the following, which involves the learners reading the sentences aloud, while maintaining the stress on the content words as more function words are added:

	●			●			●	
	MICE			EAT			CHEESE	
The	MICE			EAT			CHEESE	
The	MICE			EAT	the		CHEESE	
The	MICE	will		EAT	the		CHEESE	(Celce-Murcia, 2010, p.215)

This is followed by suggestions for communicative practice activities focusing on the target features, such as a late-night talk show role play (p.217).

It is to be assumed that research which cites *Teaching Pronunciation* in their methodology have used this procedure and that it has proven successful.

3.3.3.2 Corrective Feedback

There is also evidence that corrective feedback (CF) is an important feature of successful PI. Ellis (2006) defines CF as “responses to learner utterances containing an error” (p.28). CF can take various forms: it can be immediately after the error or delayed, and it can be implicit or explicit. An example of implicit CF would be a *recast*, where the teacher reformulates the learner’s incorrect utterance into a correct form, while explicit CF could involve a *prompt*, such as *paralinguistic signals* (such as hand gestures to identify the error) or a clarification request in order to prompt a rephrase. Alternatively, CF could take the form of the explicit identification and correction of the mistake, possibly with *metalinguistic feedback* - an explanation of the mistake. Under the term CF could also be included the more general strategy of *scaffolding*, defined as providing CF to the learner, such as recasts or form-focused prompts, which is slowly withdrawn over time, in order to allow the learner to gradually take control of their own learning (Van de Pol et al., 2010).

Research into the effectiveness of CF have generally elicited positive results in the field of PI. Saito & Lyster (2012) studied the effectiveness of recasts in a study involving 65 NS Japanese learners of English who were divided into two groups, both of which received FFI in the perception and production of the English alveolar approximant /ɹ/, but only one group received CF in the form of pronunciation-focused recasts by the teacher. Participants produced relatively controlled speech tokens from a word-reading task and a sentence-reading task, plus a more spontaneous token via a picture-description task. Results were analysed in two ways: by rater judgements of 5 NSs of English and by acoustic analysis of frequency values for formants 1, 2 and 3, plus F3 duration. Results showed that the

FFI+CF group outperformed the FFI-only group, leading the authors to propose that “FFI in conjunction with CF might be an effective and efficient way to promote not only L2 grammar but also L2 pronunciation development” (p.26). Similarly, in a study by Sardegna & McGregor (2013), scaffolding is cited as one of the techniques found to help learners improve their accuracy scores for “vowel reduction, linking, primary stress, and intonation in pre- and post-read-aloud tests” (p.182). Gooch et al. (2016) proposed that different types of CF might achieve different results. In their study of 22 L1 Korean speaking learners of English, they compared the effectiveness of FFI with recasts versus FFI with prompts in the teaching English /ɪ/. Lessons were video-recorded and production judged by 5 English NSs in a spontaneous (picture description) and a controlled task (production of 8 target words with word initial /ɪ/). Results suggested that both forms of CF were effective, but that prompts improved intelligibility, while recasts helped participants improve accuracy (nativeness) in their production. Similarly, a meta-analysis of CF studies in PI by Saito et al. (2021), warns that there is ‘individual variability’ (p.13) in the effectiveness of different strategies, but concludes that results tend to show that CF promotes accuracy, both at the segmental and suprasegmental level.

In summary, there is general agreement that CF can be effective in PI, but that more work needs to be done to identify how different strategies might adopted depending on the goal of instruction.

3.3.3.3 CAPT (Computer Assisted Pronunciation Training)

There has also been considerable research into the use of CAPT in PI. Thomson & Derwing (2015b) suggested that, at the time of writing, 59% of CAPT studies focused on segmentals, but there are now a growing number of online resources and mobile applications with a wider focus. Clearly not all of them are built on reliable research. Lippi-Green (2012) notes that anyone with access to the internet can create an accent reduction website, which she compares to individuals who “have developed a miracle diet and charge money for it” (p.229). Similarly, O’Brien (2021) notes that accent-focused apps frequently focus on segmentals and warns that such a focus may not play a role in comprehensibility. However, there are many well researched online pronunciation tools - see Henrichsen (2019) for a useful overview. These can involve perception training (*Youghlish*, *Sounds of English*, *BBC Pronunciation Workshop*, *British Council Sounds Right*), or production training (*Pronunciator*, *Sounds of Speech*, *Mango Languages*) as well as more general language learning resources such as *Duolingo* and *Busuu*. For example, Fouz-González (2019) reports success in a 3-week segmental training programme for L1 Spanish learners of English using podcasts, more specifically the BBC *Six Minute English* podcast. Participants were first given explicit instruction regarding the place and manner of articulation of the target sounds, which were the /s/ and /z/ contrast and the pronunciation of /b/, /d/ and

/g/ stops in intervocalic position. Participants then listened to one of the podcasts and completed awareness-raising activities. This was followed with a recorded read-aloud task, where target features were highlighted for the listeners and recordings were then subjected to peer feedback. The study reported a positive impact of this intervention on the participants' perception and production of the target sounds.

The general consensus to date is that these tools may be a useful addition to the intervention repertoire, but they do not replace the teacher. Yoshida (2018) comments that "technology can be a valuable tool in teaching and learning pronunciation, but it is not a panacea. There are limitations to what any tool can do and how it can be used" (p.209). Meanwhile Lee et al. (2015), in their meta-analysis, state that, "studies that provided PI using technology, whether entirely or in part, produced smaller effects than those that relied exclusively on human-delivered instruction" (p.361). Thus, research seems to favour the judicious use of technology in PI and that CAPT should not replace 'human-delivered instruction'. While the recent introduction of AI will undoubtedly influence this discussion, this is beyond the scope of this research.

3.3.3.4 Contrastive Analysis

Saito (2012), in his review of 15 PI studies, states that studies which factored in L1 background when selecting target features for instruction tended to show success. This is a finding supported in the Lee et al. (2015) meta-analysis, where they state that "the results of our study support an approach that treats sets of features that align with learners' needs, backgrounds and first languages" (p.361). Research which focuses on predicted L1 interference problem areas for speakers of EFL has shown positive results, for example, European Portuguese (Rato, 2014), Brazilian Portuguese (Alves & Magro, *ibid.*) Spanish (Fouz-González, *ibid.*), Slovak (Kralova et al., 2020) and Vietnamese (Nu, 2015).

The Kralova et al. (2020) study, compared the production of English short vowels /ɪ/, /e/, /æ/, /ʌ/, /ɒ/, /ʊ/ and /ə/ by Slovak learners before and after 10 weeks of phonetic training. Participants were 80 EFL university students whose L1 was Slovak and who were evaluated as having B1-B2 level of competence in English. They were divided into 2 groups: the experimental group, which received Slovak-English contrastive phonetic training and the control, which received training solely on the phonetic system of English. The analysis involved the pre- and post-test spectral analysis of the 7 short English vowels using the Speech Analyser system (version 2.7), as compared to production of 5 British NSs. Results showed that both groups showed improvement pre- to post-test but that, in the post-test

condition, the experimental group outperformed the control group suggesting that the contrastive analysis approach was more effective than the English-only phonetic training. Similarly, Nu (2015), conducted a study involving 8 L1 Vietnamese adult EFL learners, in which she explicitly introduced the participants, over a 20-hour intervention, to sounds which had been identified as problematic for L1 Vietnamese speakers, particularly the fricatives /ð/ and /θ/ and affricates /dʒ/ and /tʃ/, which do not exist in Vietnamese. She reported using a variety of methods including the phonemic chart, videos and sagittal section diagrams to teach articulation and various listening, repeating, matching and transcribing exercises, along with games and roleplays to practise the problem phonemes. Findings suggested that students had become more aware of the distinction between the sounds, but reported that they “still make mistakes occasionally” (p.32) in speech.

However, some research has questioned the power of contrastive analysis in PI. Cai and Lee (2015) compared the sound systems of Mandarin Chinese and Thai and analysed the L1 influence on the L2 Mandarin Chinese of 5 L1 Thai speakers who were learning Mandarin Chinese and were at various levels of proficiency. Using a read-aloud paradigm, the speech tokens were evaluated by the authors and by a Chinese teacher at a Chinese university and it was noted that “some differences between L1 and L2 do not cause difficulty in L2 learning, while some similarities do” (p.719) and that, therefore, contrastive analysis is not particularly effective in predicting pronunciation needs. This was also the conclusion of Munro (2018), mentioned in Section 2.2 above, whose synthesis of empirical research in the field of pronunciation error prediction, highlights the fact that a reliance on contrastive analysis between the L1 and L2 may be misleading since it fails to take nonlinguistic factors such as experience, aptitude and motivation into account.

In short, while some researchers have highlighted the benefits of contrastive analysis in PI, as previously argued in Section 2.2, L1 interference is not the whole story in predicting L2 pronunciation attainment and should therefore not be the only factor guiding PI.

3.3.3.5 Oral Reading

A further technique for working on pronunciation modification is proposed by Ricard (1986), who describes a PI course, using a technique she calls *oral reading*, which involves the use of read-aloud passages to enable the learner to focus on aspects of pronunciation, without having to worry about the content of their utterance. Ricard did not provide any research data regarding the effectiveness of this approach, but research has suggested that it might be effective in identifying learner pronunciation needs (Levis & Barriuso, 2012; Shak et al, 2016) and to work on segmentals features (Bratton, 2019) in

learners who are competent readers in the L2. In the Bratton (2019) study, read aloud tasks, plus FFI which focused on the / i:/ /ɪ/ distinction were effective in modifying the speech of an L1 Mandarin learner of English.

3.3.4 Summary - Structure-Focused Pronunciation Instruction

In conclusion, production training has shown success in pronunciation modification studies. A focus on both segmental and suprasegmental features, which includes a communicative component, would appear to be the most effective intervention. In terms of the content of intervention, a broadly PPP approach, starting with discrimination tasks, minimal-pair drill, reading of short passages and sentences out loud, building to free practice as suggested in *Teaching Pronunciation*, has shown promise. Corrective feedback, content based on contrastive analysis, the judicious use of CAPT and oral reading have also provided positive results in PI research studies.

While PI has generally been shown to be effective in achieving accent modification, studies tend not to test for transfer to spontaneous speech production. A few of the studies cited above found transfer to spontaneous production in the form of picture story narrative tasks (Derwing et al., *ibid.*; Gordon & Darcy, 2019) but much of the research does not include a spontaneous speech production component (Missaglia, *ibid.*; Hahn, *ibid.*; Saito, 2013; Gordon & Darcy, 2016). Thomson and Derwing (2015a) in their review of 75 L2 pronunciation studies mentioned above, note that only 20% of studies included spontaneous speaking tasks and that, within that group, some of them allowed time for preparation, which means calling them ‘spontaneous’ is questionable.

Even if there is a clear link between focused PI and spontaneous speech production, there are those who question whether such an approach is productive. Gilbert (2014) provides this sanguine metaphor:

“Learning the pronunciation of a second language is something like learning to play tennis. If you concentrate too much on trying to remember what to do with your wrist and, at the same time, try to remember how to position your feet, shoulders, and knees while also trying to remember to keep your eye on the ball, the combination of too many things to think about tends to keep you from putting everything together in a flowing movement” (p.128).

The suggestion here is that a conscious focus on pronunciation is extremely problematic during online speech production, since working memory is taken up with the construction and production of the message. More work needs to be done to establish whether production gains at the segmental and suprasegmental level described above can transfer to spontaneous speech. Meanwhile, research is

beginning to suggest that intervention might be more effective if students are encouraged to go beyond a focus on specific features of pronunciation, be they segmental or suprasegmental, and to think more globally about PI in terms of helping the learners “to develop awareness of themselves as speakers ...” (MacDonald, 2018, p.249). What this might entail is the focus of the following section.

3.4 TOP-DOWN INSTRUCTION - BEYOND THE SUPRASEGMENTAL

LaScotte et al. (2021) describe how the research and teaching of L2 PI largely falls into two groups: the ‘bottom-up’ approach, which is essentially structural, focusing on de-contextualised phonological forms, and the ‘top-down’ approach in which communication is paramount. As such, the PI approaches described in Section 3.3.3 above can be described as bottom-up since the pronunciation component focuses on drawing the learners’ attention to specific features of speech. While ‘top-down’ has sometimes been equated with suprasegmental in PI (see, for example, Pennington, 1989; Dalton & Seidlhofer, 1998), it is becoming clear that PI needs to go beyond the suprasegmental to also include wider issues which may be pertinent to pronunciation attainment. Pennington (2019a) describes how, in this revised view, “a top-down pronunciation curriculum would start at the level of broad social context, suprasegmentals, and related behaviours, then systematically progress ‘downwards’ to the segmental level” (p.381). LaScotte et al. (ibid.) point out that the more structure-led ‘bottom-up’ approach to L2 PI research “treated the influence of social and communicative context as negligible” (p.145). In the new ‘top-down’ paradigm, while segmental and suprasegmental instruction are still important elements of PI, they should only be addressed once the learner has been asked to think about wider issues influencing pronunciation, such as individual pronunciation goals (Section 1.5.2), pronunciation self-awareness (Section 2.4.1.4) and the influence of identity (Section 2.7), which provoke the learner to think about their pronunciation and motivates them to make informed choices regarding goals. Instruction then focuses on what Thornbury (1993) describes as a ‘top-down holistic paradigm’ (p.126) which focuses on strategies, such as mimicry, shadowing, the ‘golden sentence’ and voice quality work to prioritise work on fluent speech, rather than focusing on specific structures. The following sections look at ‘top-down’ strategies aimed at helping the learners to become aware of themselves as speakers and the broader social context as it pertains to L2 accent.

3.4.1 Student-centred Learning

Bernaus (2010) comments, “if learners feel that they are their own masters in their learning they will be motivated to learn and of course to attain higher levels of achievement” (p.185). In order to be in a position to make realistic choices with respect to accent, learners need to be aware of the difference between intelligibility/comprehensibility and accentedness, and to be aware of the debate around the target accent of instruction, so that they are in a position to make informed choices about their own pronunciation goals. Furthermore, they need to have a basic idea about the overall findings of pronunciation research. For example, late learners often report believing that accent modification is just not possible if learning starts after a certain age (Flege & Mackay, 2011; Moyer, 2017; Muñoz & Singleton, 2011; Piller, 2002), but, as noted in Section 2.1, research would suggest that this is not the whole story (Birdsong, 2007; Dollman et al., 2020; Moyer, 2021). Meanwhile, most learners are unaware of the influence of ID variables such as personality and identity on accent. Therefore, PI needs to include a basic overview of these factors and, where possible, offer strategies to offset negative influences. Miller (2001) says the following: “although students are ultimately responsible for their own improvement, teachers are responsible for providing them with the concepts and tools for controlling their progress” (p.183). The following section is an overview of such ‘concepts and tools’ which have shown success in research studies.

3.4.2 Mimicry/Shadowing/Echoing/Mirroring/PPR

In Section 2.4.2 above, it was suggested that mimicry ability was considered a feature of language aptitude and that the ability to mimic speech has been shown to lead to positive L2 pronunciation acquisition outcomes. Kendrick (1997) found that, for the two participants who scored highest for pronunciation accuracy in her study, the two mimicry tests proved to be the most reliable predictors. She concluded that, along with instruction and talkativeness, aptitude for mimicry “was another variable which had significant correlations with success in pronunciation” (p.556). There is some evidence that learners can be helped to improve their ability to mimic.

One strategy which has been used with some success is *shadowing*, described by Luo et al. (2008) as an intervention where learners imitate a speech sample “as closely and quickly as possible” (p. 35) after hearing it. Foote & McDonough (2017) report a study where participants practised shadowing short NS dialogues for 8 weeks and were rated pre-, mid-, and post-test for comprehensibility, accentedness and fluency by 22 NS raters. Results showed that participants improved in all areas over the course of the intervention, with the exception of accentedness. Similarly, Martinsen et al. (2017) were able to

show that shadowing short video excerpts by NS English learners of French resulted in French NS judges identifying an improvement in intelligibility at the end of a 10-week training period. Criticism of the strategy as a return to audiolingual methods and “word-for-word repetition, in the same language, parrot-style” (Lambert, 1988, p.266), is countered by Foote and McDonough’s comment that shadowing “should not be considered as a replacement for classroom-based pronunciation instruction” (ibid., p.50). Rather than being a methodology in itself, ‘shadowing’ is just one tool available to the learners in their accent modification journey. Furthermore, Ding et al. (2019) point to the need for students to make their own choices regarding their shadowing target and suggest that “second-language learners are more likely to succeed when they imitate a speaker with a voice similar to their own, a so-called “golden speaker” (p. 51). The ‘golden speaker’ refers to research by Probst et al. (2002) which indicated that L2 learners who imitated a voice similar to their own improved more than those who imitated a poorly matched voice.

While shadowing involves the immediate repetition of target input. Chung (2017) describes the *Echo Method*, which uses what she calls *echoic memory*, but is better known as the *phonological loop* (Baddeley & Hitch, 1974), the short-term memory responsible for the temporary storage and manipulation of verbal and auditory information estimated to last a few seconds. She proposes listening to a target phrase and then allowing the brain to “absorb the ‘echo’ of a spoken phrase” (p.141) before imitating it and suggests that allowing this time to absorb the phrase makes the Echo Method more effective than Shadowing but does not offer research evidence for this claim.

Mirroring takes the mimicry idea a step further. Meyers (2018) describes the Mirroring Project in which ITA students at Minnesota University chose a short video of a ‘model’ speaker, whose speech they analysed, along with paralinguistic features such as gestures and facial expressions, in order to completely ‘mirror’ their target. Learners then produced their own trial short video of the mirrored speech and received feedback on their performances from both teacher and students to help refine their performance in preparation for a final video. The author describes how students initially focus on the “big picture” (p.1), the overall impression created by the speaker, and gradually work downwards to smaller details of performance. The mirroring technique differs from shadowing in that it is not simultaneous and the focus is on non-verbal as well as verbal behaviour. Tarone & Meyers (2018) report of the gains, particularly in primary stress and pitch range, made by one student using this strategy.

Rather than the repetition of stretches of speech, the imitation of a single sentence, a *template*, spoken by a target speaker and dubbed by the author, the *golden sentence* has been proposed. Gilbert (2014) discusses the use of such a template in PI in the SLA classroom, which she describes as “probably the

most useful reference tool a student can possess” (p.130). She describes a process in which the learner listens many times to the template, before producing and refining the template through repetition, peer feedback and sentence analysis. The aim is to give the learners “an opportunity to absorb into their personal long-term memory banks an accurate sample of spoken English in which all the levels of the prosodic system are present” (p.130). At present, there appears to be a lack of empirical evidence into the effectiveness of the ‘golden sentence’ technique, but evidence from shadowing and mirroring described above suggests that this might also be a useful PI strategy.

In Section 3.4.4.5 above, it was suggested that *oral reading*, the use of read-aloud passages to work on specific structures, has shown promise as a pronunciation diagnostic tool and for working on segmentals. There is also research suggesting that reading could be used as a strategy to work on pronunciation in general, rather than as a tool for focusing on specific structures. Luo (2016) reports on a strategy she calls *peer-reviewed pronunciation reading* (PPR) in which she used CAPT to enable NS Taiwanese learners of English to record a short reading passage and compare it with the recording of an NS in order to help them modify their L2 English pronunciation. Learners also received peer and teacher feedback on their recordings. Results showed that learners who had taken part in the intervention achieved superior outcomes, as judged by experienced raters, to those who had only received English instruction. Similar positive effects for oral reading with feedback are reported by Adrián (2014), who achieved pronunciation gains in learners who took part in a programme using recorded read-aloud tasks with feedback. Oral reading used in this way is similar to the other mimicry strategies described above, but using the reading passage as scaffolding.

3.4.3 Channeling/Roleplay

Levis & McCrocklin (2018), describing features of exceptional learners (see Section 2.1 above) comment that they are, “socially outgoing in how they approach learning and use language in as many ways as they can” (p.78) and this agrees with work on personality in language aptitude discussed in Section 2.4.5. While it is unlikely that instruction can, or should, for example, help a shy student to become socially outgoing, it has been suggested that there are strategies which will help the learner overcome certain personality traits which might be holding them back from achieving the desired pronunciation outcomes.

Moreno (2016) showed how a Spanish speaker who started learning English in his teens was able to *channel* the voices of NSs he had met and, in the process, was able to achieve, what the author described as, more native-like pronunciation than when he was speaking with his own L2 voice. This

channeling strategy goes beyond the mimicry strategies of Section 3.4.3 above since participants are not just copying a target, but are invited to pretend they *are* the target. Channeling was used in a study by LaScotte & Tarone (2019) involving 10 ESL speakers of different L2 language abilities and resulted in the authors concluding that “learners enacted imagined ‘voices’ that were significantly more accurate and fluent compared to their narrative baseline voices” (p.95). The finding was backed up by a subsequent study by the authors (LaScotte & Tarone, 2022). Pennington (2019a) takes a similar approach when she suggests that pronunciation modification involves altering aspect of the self, psychological, social and cultural, and suggests that, through the use of roleplay activities, “the L2 speaker can alter their pronunciation when they are in a sense *being someone else*” (p.377). Tasks involving ‘channeling’ and roleplays might help the learner to jettison aspects of personality and identity which may, consciously or subconsciously, be holding them back from realising their accent goals as discussed in Sections 2.6 and 2.7 above. This was also the finding of Pennington and Waxler (2017), who reported that openness and empathy, personality traits addressed in Section 2.6 above, can be fostered when learners relate to a fictional character, such as performing a character in a roleplay. That drama activities can promote pronunciation accuracy has been reported in other research findings (Korkut & Çelik, 2021; Florare Bora, 2021).

3.4.4 Voice Quality/Articulatory Settings

In Section 1.1 above, it was discussed how *voice quality*, sometimes referred to as *voice settings* or *articulatory settings*, have been proposed as a level of pronunciation above suprasegmental features (Jones & Evans, 1995, Pennington & Richards, 1986; Rogerson-Revell, 2011) Pennington and Richards (1986) describe *voice-setting features* as “the tendency of speakers of a particular language to adopt certain habitual positions of articulation in connected speech, resulting in a characteristic voice quality” (p.209). Laver (1980) describes *voice quality* as being the supralaryngeal setting (above the larynx) and overall muscular tension in the articulators which determine whether we evaluate a voice as whispery, creaky, nasal, falsetto and so on. Jones and Evans (1995) noted that different languages have different ‘voice quality settings’ which go beyond prosody to influence the overall impression of the language’s auditory character. They describe, for example, how French has a ‘labialized and dentalized quality, while Mandarin Chinese has a nasalized, tense voice quality” (p.245). Todaka (1993) was able to show that, in a study of Japanese/English bilinguals, both men and women spoke with a higher pitch when speaking Japanese compared to speaking English. Rogerson-Revell (2011, p.38) refers to the work of Honikman (1964), who identified the following features of English speakers: general lack of

muscularity (laxness in English consonants), limited lip movement, a preference for centralised vowels (/ə/), many alveolar articulations (t/d/l/n/s/z) and a frequent return to rest position within articulations.

While voice quality has been identified as an important dimension of production (Pennington & Richards, 1986), Jones & Evans (ibid.) comment that “little attention has been paid to it in teaching materials”. They argued that giving learners an overview of English settings would “sensitize them to differences in voice quality between English and their native language” (p.250), but did not provide any research evidence for this claim. Thornbury (1993) offered similar advice saying that “until the learner is able to approximate the voice-setting features of the target language, work on individual phonemes is largely whistling in the dark” and proposed that by using “a holistic, top-down paradigm ... the bits might take care of themselves” (p.128). He proposed a series of possible strategies for raising student awareness to voice settings such as watching videos with the sound switched off to see if students could identify the language being spoken on the basis visual information only, the use of cognates (e.g. English *sardines*/Spanish *sardinas*), perception exercises to identify voice quality features and pair/group roleplays using exaggerated features of the accent being studied. However, Thornbury offers no research findings into the effectiveness of these strategies.

Rogerson-Revell (ibid.) similarly offers activities such as writing phrases in the L1 and L2 and articulating them without sound to see if peers can guess the language, and speaking in the L1, but with an English accent, in order to try to identify the L1/L2 differences in articulatory settings. However, as Celce-Murcia et al. (ibid.) point out, “we know that voice-quality differences do contribute to a foreign accent” (p.32), but evidence into its effectiveness in PI is sparse and “more research needs to be done” (p.32).

3.4.5 Summary – Top-Down Pronunciation Instruction

In summary, there is mounting evidence that drawing learner attention to their own production, alerting them to psycho-social factors influencing their production and offering strategies to help them modify their production are effective at provoking accent modification. Sardegna (2022) says that the job of a teacher is to empower the learner to continue learning once they finish the course. If the teacher has done their job well, according to Sardegna, they will have gained explicit pronunciation knowledge, a range of resources and strategies and an awareness of their learning goals, which can support them as they proceed on their “pronunciation learning journey” (p.375). The following section looks at the degree to which these findings, and the findings proposed in the sections on Perception and Production Training have been integrated into mainstream SLA instruction.

3.5 CURRENT ATTITUDES TO THE TEACHING OF PRONUNCIATION

Given the robust research findings into the success of PI, perception and production training and the use of strategies which help the learners gain an awareness of themselves as L2 speakers, are these findings being translated into classroom practice? It would appear not, and the reason seems to be largely due to a lack of appropriate teaching materials and lack of PI training for teachers.

As recently as 2018, Darcy bemoaned the fact that, “despite being widely recognized as important, pronunciation does not receive equal attention in the language curriculum” (p.13) and this is certainly true for general ESL/EFL coursebooks. For example, a stalwart of the English language class, the *Cutting Edge* series from Longman Pearson, had no explicit pronunciation work prior to the third edition in 2016. Recent editions now include a small pronunciation box for each chapter, focusing on one stand-alone feature of pronunciation with controlled practice exercises. For example, *Cutting Edge Intermediate Third Edition* (Cunningham et al., 2021), Chapter 1 covers the use of strong and weak forms of auxiliary verbs with a brief explanation followed by a listen and repeat exercise. There are signs that pronunciation is beginning to receive more attention. In the latest edition of *Speakout Pre-Intermediate* (Clare & Wilson, 2022) from the much-used *Speakout* series from Pearson, a respected EFL publisher, produced in conjunction with the BBC, PI focuses on discrete elements of pronunciation, but there are several items in each chapter and, although the treatment is generally an explanation followed by controlled practice, there are also exercises to integrate learned features into less controlled practice. For example, in Chapter 1, an exercise on linking (assimilation) ends with a task encouraging learners to use linking in open-ended ‘*how often do you ...?*’ questions (p.21).

However, despite the success of research into PI strategies which go ‘beyond the suprasegmental’ and address pronunciation from a top-down approach, this has not filtered down to mainstream EFL materials. While well-researched publications such as *Teaching Pronunciation* by Celce-Murica et al. (ibid.) and Rogerson-Revell’s (2011) *English Phonology and Pronunciation Teaching* include sections on social and psychological influences on accent and strategies to promote learning, these are complex and detailed publications aimed at the experienced teacher with a knowledge of phonetics and phonology. Mainstream pronunciation classics, such as the *Headway* pronunciation series (Oxford University Press) and the *English Pronunciation in Use* series (Cambridge University Press), do not go beyond suprasegmental training.

Derwing & Munro (2009) suggest that a lack of attention to pronunciation means that teachers are left to “rely on their own intuitions with little direction” (p.379). Foote et al. (2011), in a survey of 67 language teachers in Canada found that 75% of respondents indicated that they would like more pronunciation training, with comments such as: “anything would be nice!” and “too many teachers

avoid teaching pronunciation because they lack confidence in their own ability to succeed with it” (p.16). Foote et al. (ibid.) found that, of those teachers who reported working on pronunciation, most focused on a single aspect of pronunciation (largely individual segments), which, as discussed above, may not have had a significant effect on outcomes. Lack of teacher training in pronunciation has also been noted by other researchers (Baker, 2014; Murphy, 2014; Darcy, 2018).

This lack of adequate training in pronunciation is combined with a perception by teachers that pronunciation instruction is somehow a specialist area. Celce-Murcia et al. (2010), in the 2nd edition of their highly influential training manual, *Teaching Pronunciation* mentioned previously, state that “to be adequately prepared to teach pronunciation, teachers must have at their disposal a working knowledge of articulatory phonetics, theories of second-language phonological acquisition, and an up-to-date command of techniques and procedures to use in the classroom” (p.35). Similarly, Thomson & Derwing (2015a) call for “skilled instructors who can determine the pronunciation needs of a given learner” (p.14). In a perfect world, this is undoubtedly the best option, but it is unrealistic. The gold standard qualification for Teaching English as a Foreign Language (TEFL), the Cambridge Certificate of Teaching English to Speakers of Other Languages (CELTA), is a one-month English language teaching course whose website claims that “three out of four English language teaching jobs require a CELTA qualification”. To be accepted on the course, NSs require no specific qualifications, but applicants are required to complete a brief grammar test, the answers to which could be easily found online. It is impractical to believe that, in a one-month course, it is possible to offer novice teachers, who may have only basic language qualifications, anything other than the most cursory introduction to language instruction in general, let alone ‘a working knowledge of articulatory phonetics’.

It is extremely important that if pronunciation teaching is to be (re-)integrated into the L2 curriculum, the methodology has to be accessible to teachers without an expert knowledge of phonology. Levis & Wu (2018) recognise this need when they note, “for teachers who do not usually teach pronunciation, materials not only have to be available but also understandable in how to use them effectively and in what they are meant to achieve” (p.7). So, L2 teachers need to have at least some training in PI, along with easy-to-use PI materials which reflect current research findings in phonological acquisition. This does not appear to be the case, at present.

3.6 SUMMARY – INFLUENCE OF INSTRUCTION ON L2 PRONUNCIATION ATTAINMENT

In this chapter, it has been proposed that the first step in any approach to PI should be work on phonological self-awareness (Couper, 2015; McGregor & Reed, 2018) and that there are effective strategies for raising self-awareness such as the recording of and listening to the learner's own speech samples (Couper, 2015; MacDonald, 2018), self-report questionnaires (Sardegna & McGregor, 2013) and listener feedback (Miller & Ng, 1994; Tsunemoto et al., 2022).

It has also been suggested that L2 PI falls, broadly speaking, into two approaches (LaScotte et al., 2021): the 'bottom-up' approach, which is essentially structural, focusing on de-contextualised phonological forms, and the strategy-led 'top-down' approach in which communication is paramount. Research into the effectiveness of PI in promoting pronunciation modification have generally shown positive results for both the structure-led (Saito, 2012; Lee et al. 2015) and the strategy-led approach (Chung, 2017; Foote & McDonough, 2017; LaScotte & Tarone, 2019); and yet only the former seems to influence current pronunciation materials for classroom use.

The next chapter details a study in which these two important PI issues are the focus: the effectiveness of strategies to raise phonological self-awareness and the relative effectiveness of a bottom-up versus a top-down approach.

4. RESEARCH METHODOLOGY

“The learner is not only an object of scientific curiosity,
but also a crucial witness of his or her own learning process” (Dewaele, 2005, p.369)

In the previous chapter, it was proposed that various types of intervention have proven effective in achieving accent modification: perception training, structure-led bottom-up production training and strategy-led top-down instruction. Building on that research, this chapter explains the motivation behind this research and the resulting research questions. There follows a detailed justification for the choices regarding methodology: how raters were chosen, the type of speech tokens collected, how tokens were rated and how participants were chosen. This is followed by a detailed description of the programme followed by the two intervention groups and an explanation of how and when data was collected.

4.1 RESEARCH QUESTIONS

In Section 2.4.1.4 above, it was suggested that learners might struggle to evaluate their own accent and show a tendency to over-estimate their performance in L2 speech (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al., 2016), while good self-monitoring skills appear to be a feature of exceptional learners (Moyer 2014a; Moyer, 2018; Muñoz & Singleton, 2011). In Section 3.2.2, it was argued that an awareness of the learner's current production needs to be one of the first steps of any intervention to modify pronunciation (Couper, 2015; McGregor & Reed, 2018; Sardegna & McGregor, 2013) and that research has shown success in helping learner to gain self-awareness (Lappin-Fortin & Rye, 2014; MacDonald, 2018; Tsunemoto et al., 2022). This study therefore hoped to build on this research by comparing learner pronunciation self-awareness pre-intervention with self-awareness following an intervention which included awareness raising activities: recording and listening to a passage of their own speech (Couper, *ibid.*; McDonald, *ibid.*) and receiving teacher and peer feedback (Tsunemoto et al. *ibid.*). It was expected that receiving pronunciation feedback from the instructor and other learners in the group, along with encouraging speakers to record and listen to their own output, would give participants greater awareness of their current accent.

Secondly, in Section 3.1 above, it was proposed that exposure alone may be insufficient to promote L2 pronunciation acquisition in late learners (Schmidt, 2012), but Section 3.3.3 discussed how PI has generally proved effective in provoking change (Kennedy & Trofimovich, Lee et al., 2015; Saito, 2012; Thomson & Derwing, 2015a). In Section 3.4 above, it was described how L2 PI largely falls into two

groups: the structure-based bottom-up approach and the strategy-based top-down approach (LaScotte et al., 2021). Research into global strategies, those which focus on overall performance rather than specific structural features, have shown promising results, such as shadowing (Foote & McDonough, 2017; Martinsen et al., 2017), echoing (Chung, 2017), mirroring (Meyers, 2018), channeling (La Scotte & Tarone, 2019; Moreno, 2016) and roleplay (Pennington & Waxler, 2017). Advocates of a top-down approach have also highlighted a need to raise student awareness to wider issues influencing L2 pronunciation such as identity (Pennington, 2019), social context (LaScotte et al., *ibid.*; Pennington, *ibid.*) and a need to approximate voice settings (Rogerson-Revell, 2011; Thornbury, 1993). Research has been done to compare the relative effectiveness of a bottom-up versus a top-down approach in the area of SLA listening (Orii-Akita, 2014; Yeldam, 2016) and vocabulary learning (Lee & Yoon, 2019; Moskovsky et al., 2015), but as yet there appears to be no research comparing a bottom-up versus a top-down approach in a general PI programme. The current study hoped to begin to address this lacuna by comparing the two approaches. It was predicted that the top-down approach, with its more holistic and communicative-led focus, would produce superior results to a purely structural syllabus in line with the recommendations of prominent researchers (LaScotte et al, *ibid.*; Moyer, 2018; Pennington, *ibid.*).

In order to attempt to shed light on these two areas: learner self-awareness and the bottom-up/top-down debate, the research questions are as follows:

1. To what extent do learners' pre-intervention assessments of their own spontaneous speech production correspond to those of NS raters?

Prediction: In line with research cited above (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al., 2016), it was predicted that many learners will judge their accents more favourably than will NS listeners.

2. Will either or both of the interventions to influence pronunciation produce superior results to mere exposure to a native speaker accent in late learners?

Prediction: In agreement with previous research (Saito, 2012; Lee et al., 2015; Kennedy & Trofimovich, 2017; Thomson & Derwing, 2015a), pronunciation intervention will prove more effective than mere exposure.

3. Will a top-down strategy-led approach bring late learners closer to achieving a native-like accent than a bottom-up structure-led approach?

Prediction: Top-down strategies to work on pronunciation will be more effective than bottom-up at modifying spontaneous speech in line with the recommendations of prominent researchers (LaScotte et al, *ibid.*; Moyer, 2018; Pennington, *ibid.*).

4. Can intervention influence the learners' ability to evaluate their own spontaneous speech production to correspond to that of NS raters?

Prediction: In accordance with previous research which has shown success in helping learners to gain self-awareness (Couper, *ibid.*; Lappin-Fortin & Rye, *ibid.*; MacDonald, 2018; McGregor & Reed, *ibid.*; Sardegna & McGregor, *ibid.*; Tsunemoto et al., 2022) intervention will help to raise learner phonological self-awareness.

In an attempt to answer these questions, the following sections provide a detailed description of the methodology used and the rationale behind decisions made.

4.2 EVALUATING PRONUNCIATION

In 2004, Jesney bemoaned the fact that “the elements that contribute to the perception of foreign accent, and, indeed, the objective characteristics of it, remain ill defined” and that “the precise methodology utilized in making such foreign accentedness ratings remains unstandardized ...” (p.2). In the ensuing years, while the methodology used to evaluate NNS speech tokens has remained unstandardized, much has been written on the topic, with discussions centering around *who* should judge, *what* they should be judging and *how* to judge. The following sections give an overview of findings in these three areas and a justification of subsequent decisions as they pertain to this research.

4.2.1 Who Should Judge: Choosing Raters

If a foreign accent, as has previously been suggested (Chapter 2), can best be defined in terms of subjective listener judgements, it is important to establish what sort of listeners would provide the most useful feedback for this study.

4.2.1.1 Native Speaker/Non-native Speaker Raters

In 2010, Wilkerson commented on “the tradition to privilege NS judgments in FL [foreign language] contexts” (p.144). In the past, the vast majority of studies relied on NS raters, but

this is a tradition which has been questioned by some in the field. Research into the differences between NS and NNS raters has come to mixed conclusions.

Major (2007) carried out research into ratings of Brazilian Portuguese spoken by NSs and NNSs. Listeners were NSs of Brazilian Portuguese, English NSs with knowledge of Portuguese, English NSs without knowledge of Portuguese and ESL learners without knowledge of Portuguese. Perhaps surprisingly, Major concluded that “results show striking similarities across groups, suggesting that listeners’ first and second languages do not strongly affect rating of foreign accents” (p.539).

In contrast, a growing body of research suggests otherwise. Wester and Mayo (2014) found that NNS raters in their study “judged talkers with non-native accents less harshly than did native English listeners” (p.749), which raises the possibility of slight differences in the judgements of the two rater groups. However, other studies have identified much larger differences between NS and NNS raters. Riney et al. (2005) compared accentedness ratings of English language tokens by Japanese and American raters. They noted that, “Japanese listeners used primarily nonsegmental parameters ... to make perceptual judgments, whereas segmental parameters had a relatively minor role” (p.441) but that NS listeners exhibited the opposite tendency. This led the authors to conclude that the two rater populations were perceiving the tokens in quite different ways. This is a view supported by Wilkerson (2013), who studied NNS/NS ratings of German and concluded that “NNSs might vary from NSs – and perhaps even from each other – in how they detect, rate, and describe native and non-native accent” (p.106-7). Wilkerson is saying that, not only are there differences between NNS and NS raters, but that there is a lack of inter-rater reliability among NNS raters. Clearly, that is not to say that one group is inherently ‘better’ than the other, but given the suggestion that NS raters might be better able to perceive fine-grained distinctions in tokens produced in their L1 and show greater inter-rater reliability, it was decided to use only English NS raters for this study.

An area which would appear to have attracted little or no attention in the selection of raters is the particular variety of English (British English, Australian English, ...) of the NS raters. It might be the case that, for example, an NS of British English might be less sensitive to nuances in an American English accent than would a speaker of American English. Clearly, different varieties of English show phonological variation, as well as divergences in lexico-grammar, but it has also been suggested that there might be temporal differences. Robb et al. (2004) compared American English (AE) with New Zealand English (NZE) speakers and identified a “significantly faster overall speaking rate and articulation rate for the NZE group compared to the AE group” (p.1). The lack of attention to the possibility of English variety variables

influencing rater judgements in the literature might be because much of the research has been carried out in ESL environments, that is, with NNS speakers living in the country where the target language is spoken, particularly the USA and Canada. For foreign accentedness ratings of NNSs in, for example, Canada, it would be logical to use Canadian NSs because that would be the target accent for the majority of NNSs in the country and, on a more practical level, Canadian English speakers would be the available rater population. In an EFL environment, such as the one in which this research is being conducted, the target variety of English is less clear. Since the aim of the research is to work on standard British English, it was decided to favour British English raters, so seven of the nine raters were British, with the other two being American.

Another dimension to this problem is addressed by Trofimovich and Isaacs (2012), who clarify that the raters in their research were monolingual English speakers who reported that they used English over 90% of the time on a daily basis. The researchers were obviously aware that the use of bilingual or multilingual raters, who might consider themselves NSs in more than one language, could elicit different results from monolinguals in terms of speech rating. With the increased interest in bilingual and monolingual speakers in recent years, this is an area which has begun to be researched more extensively. A study by Schmidtke (2016) suggests that “bilingual speakers are disadvantaged relative to monolingual speakers when it comes to speech understanding in noise” (p.1), which implies that there may be subtle differences between monolinguals and bilinguals in terms of speech perception. Furthermore, a study by Shintani et al. (2017) concluded that “the raters’ language backgrounds, notably the proficiency balance between multiple languages in early childhood, are related to their rating judgement” (p.849), which raises the issue that the term ‘bilingual’ might not be as homogenous as we might think and that there may be significant differences between individual bilinguals in terms of their language history and language use. In view of the issues addressed above, the raters in the current study are all NSs of English and, while some of them have a good knowledge of other languages, none of them would be considered early bilinguals.

4.2.1.2 Experienced/Naïve Raters

Another factor which may influence rating outcomes is whether raters are *experienced*, in that they have a background in SLA/Linguistics or not - so-called *inexperienced* or *naïve* raters, those with no background in the field. Thompson (1991) found that raters with no background in linguistics tended to judge accentedness more harshly than the experienced raters. It could be

that teachers and researchers are more exposed to NNS accents and therefore more forgiving of or even impervious to minor deviations in pronunciation, which might be more apparent to those whose L1 listening experience involves largely NSs.

Furthermore, Isaacs and Thomson (2013), in their comparison of 20 experienced and 20 naïve raters noted that “experienced raters are slightly more consistent in their ratings than novice [naïve] raters” (p.25). This slight bias in favour of the interrater reliability of experienced raters has been questioned in other studies. For example, Derwing et al. (2004) noted that in global, impressionistic fluency judgements to measure the temporal aspects of L2 speech, such as speech rate, fluency and pauses, inexperienced raters were comparable to instrumentally-derived linguistic analysis. This led the researchers to conclude that “rating data from even untrained listeners reflect properties inherent in the stimuli and are therefore useful in the evaluation of speech samples” (p.672).

Appel et al. (2019) proposed that, when choosing raters, it is important to go beyond a traditional reliance on experienced raters and “to target additional rater populations that may be more representative of the type of interlocutors L2 speakers are likely to encounter during real world interactions” (p.1). In the case of their research into NNS university students in Anglophone Canada, they proposed that using L1 English students as raters would more closely reflect the type of listeners with whom the NNSs were likely to interact. They even suggested that naïve raters might actually be preferential since the use of expert raters leaves open the possibility that “features being attended to simply result from the training process used to prepare raters, exposure to language teaching pedagogy and theory, or adherence to specific features listed in grading rubrics” (p.3-4). In other words, experienced raters might be making ‘unnatural’ judgements influenced by their training, which are not representative of the NS population in general. Isaacs and Thomson (2013) go so far as to suggest that it is actually preferable to employ naïve raters “when the intent is to extrapolate findings beyond the L2 classroom” (p.26).

A possible way of reducing this experienced vs naïve divide would be to provide raters with some sort of training. Many studies have used pre-rating training for raters to varying degrees, for example, Brennan and Brennan (1981a), Flege, Yeni-Komshian and Liu (1999b), Southwood and Flege (1999). However, others have specifically argued against this (e.g. McDermott, 1986), claiming that training sessions might interfere with rater judgements and therefore bias the results. In fact, the discussion may be immaterial since some research has found little difference between the judgements of experienced and naïve raters (see Bongaerts et al., 1997; Isaacs & Thomson, 2013).

In this study, since the goal of the research was to help learners in their real-world interactions, the majority of the raters (seven of the nine) had no experience in SLA or Linguistics and were not given specific training prior to the evaluations, in order to more accurately represent the speech community in which learners are likely to operate. The other two were teachers of EFL at university level.

A second aspect of this experienced vs. naïve rater debate is the level of exposure to the specific accents being rated. Jesney (*ibid.*) points to “listener-based variables, such as degree of familiarity with accented speech, that can influence the judgments made by listeners and the reliability of these” (p.2). In other words, an NS who has frequent exposure to NNS accents in general or perhaps even more significantly, to the specific accents being studied, might be more used to hearing them and therefore judge them more leniently than an NS who is unfamiliar with them. In an attempt to neutralise this possible variable, of the raters in this study, 5 live or had lived in Portugal and were therefore familiar with the accent, while 4 did not have this exposure.

4.2.1.3 Number of Raters

Past research has varied widely in the number of raters chosen to evaluate a given token. For example, Snow and Hoefnagel-Höhle (1977) used a single rater to judge the Dutch pronunciation of 47 native English speakers. In contrast, Anderson-Hsieh and Koehler (1988) used 224 raters to judge the tokens of four speakers, although each rater judged a limited number of tokens. The majority of research appears to have used numbers closer to the lower end of this continuum. Perhaps of significance in this regard is interrater reliability. If raters are, on the whole, rating in a similar manner, comparable results would be achieved even with a relatively low number of raters.

Research findings regarding interrater reliability have tended to find a degree of homogeneity between NS raters. For example, Kim and Margolis (1999) and Wemmel (2010) noted a high degree of reliability between assessors. Given the relatively high degree of interrater reliability attained in previous research, this study chose to use 10 raters for this project, which is consistent with recent research in the field (see, for example Crowther et al., 2018; Wheeler & Saito, 2022). Unfortunately, one of the raters was forced to drop out midway through the study and was therefore eliminated from this research, leaving 9 raters.

4.2.1.4 Age of Raters

While there appeared to be no reference to the influence of rater age on rating outcomes in the literature, it was decided that a range of ages would be the most appropriate choice, in order to provide a broad overview of NS perceptions of accent. Raters varied from 25 to 71 years of age, with an average age of 48.44 at the time of the initial rating.

4.2.1.5 Conclusion – Choosing Raters

A summary of the rater profiles can be found in Table 1 below. Nine monolingual, English NSs were employed as raters in this study. Seven of them were naïve raters and two were experienced. Five were familiar with Portuguese accents, while the rest were not. Seven of the raters came from a British background, two were American. In short, this research has followed to advice of Piske et al. (2001), who say that “a representative sample of raters should be recruited and not one particular type of rater” (p.195).

Table 1. Rater Profile

Rater	Nationality	Gender	Age	Familiar with EP	ESL/ Linguistics
Rater 1	British	M	27	✓	X
Rater 2	American	F	71	✓	X
Rater 3	British	F	25	✓	X
Rater 4	British	M	61	X	X
Rater 5	American	M	58	✓	X
Rater 6	British	F	47	✓	✓
Rater 7	British	F	31	X	X
Rater 8	British	M	56	X	X
Rater 9	British	F	60	X	✓

4.2.2 What to Judge: Choosing Speech Token

A second major issue to consider in speech rating research is what type of speech sample the raters should be evaluating. The speech sample produced by the speakers is generally referred to as a *token* and, as with the choice of raters, different studies have followed different paths when it comes to the length of token and the type of task should be used to elicit it.

4.2.2.1 Length of Token

The token needs to be of sufficient length in order to give the raters every opportunity to make an informed judgement. It might be logical to assume that the longer the token, the more accurate the judgement, but it would appear this is not necessarily the case. Flege (1984) found that raters were able to distinguish between NS and NNS speakers in a significant number of cases (67% of tokens) after hearing just 30 milliseconds of speech. Munro, Derwing and Burgess (2003) found that listeners performed at above chance when asked to detect NNS accents in backwards speech. While useful in establishing listeners' high sensitivity to NNS accents, such minimal tokens allow listeners only limited opportunity to make any judgements about aspects of the delivery which are affecting their judgements. To allow more nuanced judgements, longer tokens would appear to be essential.

However, evidence suggests that tokens do not have to be extensive. In a summary of the methodology used in numerous accentedness rating studies, Jesney (2004) noted that "foreign accentedness judgements can be effectively made based on tokens shorter than a sentence," but went on to say that "judges' accuracy increased to 89% when full phrases were presented to them" and concludes "longer excerpts of speech are more informative for raters" (p.4). This is a view seconded by Saito et al. (2016), who point out that, "longer speech samples may be needed ... in order to obtain a more refined picture of lexical, grammatical, and temporal correlates of comprehensibility and accentedness" (pp.18-19).

In fact, overly long speech tokens might be considered undesirable in such research. Since most of the raters are inexperienced and are merely being asked for impressionistic judgements which, as argued above, appear to be very quick to establish, longer tokens could be both unnecessary and overly taxing for listener concentration.

In view of these findings, the participants in this research were asked to produce tokens of approximately 2 minutes. These were then edited down to 30 seconds, which was thought to be long enough to allow raters to make judgements on segmental, suprasegmental and temporal pronunciation features, but short enough to maintain rater concentration levels. The editing involved, as far as possible, the removal of obvious grammatical errors and references to the speaker's mother tongue or country of origin which, it was felt, would lead to raters being influenced by these features, rather than judging purely on phonological features of the speech tokens.

4.2.2.2 ‘Range Effects’ in Speech Tokens

It is customary, in accentedness rating research, to include NS tokens along with the NNS tokens, the rationale being that if raters are unable to identify native speech, they may be considered unreliable raters. While this would seem to be a logical step, it is not without its caveats. It has been noted (see for example Flege & Fletcher, 1992) that accentedness ratings are subject to ‘range effects’. This refers to the finding that, if a large number of NS tokens are included in the sample, the NNS tokens will be judged more harshly than if there are few or no NS tokens. It would appear that raters are influenced by the overall level of the tokens and judge accordingly. In this research, in order to limit these range effects, only three NS tokens were included among the 24 NNSs tokens.

4.2.2.3 Type of Task to Elicit Speech Tokens

As mentioned above, foreign accentedness ratings remain unstandardized and this would appear to be nowhere more evident than in the task used to elicit the speech token. At one end of the spectrum, there have been very controlled tasks such as the direct repetition of NS models (e.g. Olson & Samuels, 1973; Snow & Hoefnagel-Höhle, 1977), delayed repetition tasks where there is a pause or interjection before the repetition of the NS model (for example, Piske, MacKay & Flege, 2001; Yeni-Komshian, Flege & Liu, 2000) or read-out-loud tasks (e.g., Munro & Derwing, 2001). On the other hand, some tasks have attempted to elicit more spontaneous speech tokens such as the telling of a story in response to visual prompts (e.g., Munro & Derwing, 1994; Derwing & Munro, 1997) or a free narrative in response to a written prompt (Oyama, 1976; Suter, 1976; Moyer, 1999). Crowther et al. (2018) take the idea of spontaneous speech even further and suggest that task types used thus far “do not encompass the full scope of task complexity” (p.18) that the NNS encounters in real world interactions. “What happens,” they ask, “when task goals are no longer individual, but require a speaker to take into consideration the views and actions of another?” (p.18). Their point is that, in order to be truly representative of spontaneous speech, tokens should involve the speaker in an interaction with at least one other speaker.

The more controlled elicitation tasks have been heavily criticised as being unrepresentative of natural speech (Piske et al., *ibid.*) and there is evidence that read speech tends to be more accented than spontaneous production (Oyama, *ibid.*; Thomson, 1991). However, collecting a more spontaneous token may also have its flaws. As Piske et al. (2001) point out, samples of

free speech “may contain morphosyntactic and lexical errors influencing the foreign accent rating given to non-native speakers” (p.194). This finding has been backed by other research which appears to show that it is the linguistic competence of speakers, as well as their accent, which influences NS judgements of comprehensibility and nativeness (Smith & Nelson, 1985; Crowther et al., 2018). To minimise interference from morphosyntactic and lexical competence in tokens, Crowther et al. (ibid.) conclude that “pronunciation as a specific criterion of a speaking test may be best isolated in a less complex task, where non-phonological linguistic measures are less likely to influence a rater’s assessment” (p.17).

In view of the above concerns, the following decisions have been taken with regard to the task used to elicit the speech tokens. While Crowther et al. (ibid.) are no doubt correct in their assertion that single-speaker tokens do not represent the complexity of real-world interactions, the unpredictable nature of a token involving the interaction of more than one speaker would make it extremely difficult to rate. To what extent would listener ratings be influenced by features of the interaction such as misunderstandings, disagreements and overlapping speech, to mention but a few? Highly controlled tasks are clearly not representative of spontaneous speech and therefore of dubious validity in this study, which aims to help learners to work on pronunciation in real-world situations. Even the telling of a story in response to visual prompts might be considered a contrived task in this respect. It was therefore decided to ask participants to read an open-ended question, detailed in Section 5.4 below, in order to elicit a spontaneous response. It was hoped that the general nature of the questions would elicit a relaxed response, which would more accurately represent everyday speech. In order to limit the influence of lexico-grammatical errors in the tokens, only participants with a level of English of B2 or above according to CEFR-CV (Common European Framework of Reference for Language – Companion Volume, 2018 - see Appendix A) as self-reported by participants, were accepted in the research. They, it was felt, might be expected to produce largely grammatically correct tokens. As mentioned above, tokens were, in any case, edited in order to remove, as far as possible, any lexico-grammatical errors.

4.2.2.4 Speaker Preparation Time

In the discussion of whether to use experienced or naïve raters in Section 5.2.1.2, it was mentioned that training the raters beforehand might influence their rating decisions. A similar debate exists around preparing participants for the token elicitation task. For example, in their research into English NSs’ sentence processing time of NS utterances versus Mandarin-accented

utterances, Munro and Derwing (1995b) allowed all speakers to read through the sentences silently beforehand and the Mandarin speakers were allowed to ask how to pronounce unfamiliar words and practice with one of the experimenters. The authors had valid reasons for training speakers in this way, but clearly such preparation would not be appropriate for this research, which is aiming to assess spontaneous speech production in participants. It would also be inappropriate to allow speakers to write their answer prior to recording it, as this would be subject to the same task validity issues attributed to read-out-loud tasks mentioned above. Participants in the research were therefore asked to read the question and respond immediately, to more accurately imitate real-world, spontaneous speech production. The precise instructions sent to participants can be found in Appendix B.

4.2.2.5 Ordering of Tokens

Research by Munro and Derwing (1994) appeared to show that, in a study involving two sets of speech tokens, the second set were judged more harshly, leading the authors to warn that the order in which the tokens are presented to raters may also influence judgement. Jesney (ibid.) similarly warns that, “randomization of stimuli presentation across judges is of considerable importance in order to prevent ordering effects” (p.9). To address these concerns, care was taken to randomise the order of speech tokens presented to the raters at each time point.

4.2.2.6 Appearance of Speakers

Ruben (2012) warns that judgements are affected by the appearance of the speakers and that, “listeners “hear” the pronunciation they expect to hear, sometimes with little regard to the actual properties of the acoustic signal” (p.11). In view of this possible influence of the speaker’s appearance on rater judgements, it was decided that only auditory tokens, rather than video tokens, would be presented to the raters.

4.2.2.7 Conclusions

Whatever task type is chosen to obtain the token, Crowther et al. (ibid.) have pointed out that a “possible task effect” (p.3) is almost inevitable, namely that different task types will inevitably elicit different speech tokens which, in turn, will impact research findings. While there is no perfect solution, the choice made in this study, an unprepared 2-minute speech token in

response to a written prompt which was edited down to 30-seconds, reflected the primary aim of the research, which was to capture a sample of spontaneous speech production.

4.2.3 How to Judge: Rating Speech Tokens

Having obtained speech tokens, decisions then had to be made regarding the nature of the feedback which the raters were asked to provide on each token. The following sections look at methodological issues surrounding this question and an explanation of decisions taken in this study in order to produce the Accent Rating Questionnaire (Appendix C).

4.2.3.1 Scalar versus Sliding Scale Ratings

In the field of accent research, raters have generally been asked to rate speech tokens by degrees of foreign accent, comprehensibility or intelligibility using a linear scale, such as the Likert scale, the ends of which are given titles such as ‘no foreign accent’ to ‘very heavy foreign accent’. Some researchers have questioned the ecological validity of partitioning accent ratings in such a way and suggest that the use of a sliding scale or continuum provides a greater degree of validity and granularity. Appel et al (2019), for example, used a 1,000-point sliding scale in order to target differences between comprehensibility and nativeness in listener judgements. Another type of sliding scale which has been used is the *direct magnitude estimation (DME)* technique, where listeners are asked to rate tokens in relation to each other. However, in a comparison of DME with scalar ratings of perceived accent in English of native speakers of Italian, Southwood and Flege (1999) warned that using DME might lead raters to overestimate or underestimate tokens. As Jesney (ibid.) points out, the use of DME means that “individual raters effectively devise their own scales, with the relationship between the individual numbers, rather than the raw scores themselves, being the variable of interest” (p.3). In accentedness rating, the variable of interest is proximity to a NS norm, rather than the relationship between tokens and this might explain why the majority of studies over the last decade have chosen to use scalar judgements (see, for example, Julkowska & Cebrian, 2015; Saito, Trofimovich & Isaacs, 2016; Galante & Thomson, 2017). The point would appear to be moot since, in the Southwood and Flege (1999) study referred to above, the researchers found the results almost identical whether using sliding scale judgements or scalar judgements. However, given the recent preference for scalar judgements of accent, it is this form of rating scale that has been used in this research.

Over the years, the size of the scale used in accent rating seems to have varied from a three-point scale (Tahta et al. 1981) to a 10-point scale and beyond (McDermott, 1986; Jilka, 2000). Interestingly, in the Southwood and Flege (*ibid.*) research mentioned above, the investigators used a 7-point rating scale but concluded that it lacked range and that “a nine- or 11-point scale may better measure degree of perceived foreign accent” (p.335). Similarly, Isaacs and Thomson (2013) warn that “the 5-point scales were reportedly too constraining for some raters” (p.23). Meanwhile Jesney (*ibid.*), in a comprehensive study of different rating scales used in accent assessment, comes out in support of the use of a 9- or 11-point Likert scale, stating that, due to its frequent use in recent research, “it seems reasonable to continue the use of such an approach in assessing degrees of global foreign accent” (p.3). Due to these research findings and to its frequent use in accentedness rating, it was decided to use a 9-point Likert scale in this study.

In terms of the descriptors used for these scales, these have also varied considerably. Some have used the first point on the scale to describe “no accent” (Saito et al., 2016) to mean that no foreign accent was detected. Others have articulated more specifically that ‘1’ on the scale related to “no foreign accent” (Julkowska & Cebrian, 2015) or “no non-native accent” (Kennedy & Trofimovich, 2008). On the other hand, other research has used the first point of the scale for the other end of the rating, describing ‘1’ on the scale as “heavily accented” (Trofimovich & Isaacs, 2012; Isaacs & Thomson, 2013). Thomson (2018a), in a brief summary of rating scales used in several recent studies, notes that, “while all are consistent in their use of scalar ratings ... none explains why they use slightly different descriptors, suggesting that there is no particular rationale for doing so” (p.6). It could be argued, therefore, that it is irrelevant in which direction the scale moves but, in view of a possible association between a higher number and a greater level of success, it was decided that ‘9’ on the scale would refer to the absence of a foreign accent. Furthermore, since the majority of raters in this study are naïve, the term “native speaker” was used as this was felt to be more direct than using a negative, or even a double negative (“no non-native accent”). At the other end of the scale, terms such as “strong non-native accent” (Kennedy & Trofimovich, 2008), “extremely strong accent” (O’Brien, 2014) and “heavily accented” (Saito et al., 2016) have been used. In this study, raters are being asked to judge features of accent, such as segmental and suprasegmental deviations from a native speaker norm, but also to comment on temporal, lexical and grammatical features of the token which may have influenced their judgement. It was therefore decided to use the term “not native speaker” as being the simplest term to use in opposition to “native speaker” to describe these features. For the final rating, which addresses the issue of comprehensibility, the native/not native distinction was no longer relevant so ‘1’ was defined as “impossible to understand” and ‘9’ as “very easy to understand”, in line with the standard

definition of ‘comprehensibility’ proposed by Munro and Derwing (1995b): “the listeners’ perceptions of difficulty in understanding particular utterances” (p.291) (see section 3.1).

4.2.3.2 Rating Questions

Having established how the tokens would be rated, it was then necessary to make decisions about what specific features of the token the raters were being asked to attend to. It is important to clarify from the outset that no attempt was made to conduct a more detailed linguistic analysis of the speech tokens. As Saito et al. (2017) point out, “if only linguistic analysis, rather than intuitive listener judgments, can reveal component linguistic features of L2 global constructs such as comprehensibility and accentedness, such research findings would provide limited practical applications to most pedagogical settings” (p.4). Most teachers have neither the equipment nor the training to conduct auditory analysis on their learners and since this research is aimed at informing PI in a classroom setting, a detailed linguistic analysis was felt to be inappropriate.

Question 1: Impressionistic Judgement

In view of this focus on intuitive listener judgements, the first question asks, “how close to a native speaker does the accent sound?” This is an attempt to capture the listener’s initial response to the token. Clearly, this initial judgement does nothing to explain what NSs are attending to when they make such judgements, so the following questions attempt to prompt the raters to explain more specifically what they are responding to in their global accent rating.

Question 2a: Segmental/Word-Level Features

When it comes to judging accentedness in NNS tokens, the features of speech which appear to garner the most attention are the segmental variables (as defined in Section 2). Not only is it claimed that segmentals might be an important part of listeners’ accentedness ratings, but some would argue they are the most important predictor. Saito et al. (2016) in their attempt to tease apart the constructs of ‘comprehensibility’ and ‘nativeness’ noted that “accentedness was strongly tied with pronunciation (specifically segmentals) rather than lexical and grammatical domains” (p.1). Furthermore, Hahn (2004) noted that NS listeners were able to process L2 speech more quickly if the correct primary stress was used as opposed to those with incorrect

stress or missing stressed syllables. Such lexical stress errors have frequently been linked to accent ratings (Kang, 2010; O'Brien, 2014; Saito et al., 2016). Saito et al. (2017) in their research into features attended to by listeners in comprehensibility and accentedness ratings conclude that “listeners chiefly rely on L2 pronunciation (vowel/consonant errors, word stress) when judging accentedness” (p. 14).

Given the perceived importance of segmental and word stress features identified in these reports, question 2a asks to what extent these factors influenced their judgement. The justification for merging these two features into one question was partly due to the need to keep the questionnaire short and concise, in order to maintain rater focus. However, it was also based on the argument by Pennington (1996) that “a segmental description in fact artificially breaks up the relatively continuous stream of sound that is language and arbitrarily assigns small bits of ever-shifting sound waves to individual symbols” (p.19). It was feared that naïve raters might be unable to pinpoint what specific segment of the word was provoking their judgement and, therefore, that asking them to comment on words, rather than segments, might give more accurate results.

Question 2a) asks raters to judge ‘individual sounds and words’ with the examples: ‘how ‘th’ is pronounced; how words are stressed’. The word ‘segmental’ was omitted as this was felt to be too technical for naïve raters.

Question 2b: Sentence Level Suprasegmental/Prosodic Features

Suprasegmental or prosodic features of pronunciation refer to the melodic or musical elements of speech (as defined in Section 2), such as intonation, stress, pitch, rhythm and voice quality. While these features work at the word-level, as well as in longer stretches of speech, it was felt, as previously explained, that the word/sentence distinction would be easier for naïve raters to perceive.

Kang (2010) assessed multiple features of accentedness via acoustic analysis, rather than listener judgements and concluded that “suprasegmental errors contribute as much or more to perceived accentedness than do segmental errors” (p.564). Many other studies have pointed to prosody in general (Anderson-Hsieh et al., 1992; Saito et al., 2016) or specific aspects of prosody, such as pitch range (Thomson, 2018a), reduced pitch range (Kang, 2010; O'Brien, 2014) and non-native sounding rhythm (Trofimovich & Isaacs, 2012; Thomson, *ibid.*) as having a particular relevance to accentedness ratings. However, Saito et al. (2017) point out that “it is unclear what particular linguistic training and experience is required for reliable judgements of

complex linguistic phenomena in L2 speech (e.g., vowel reduction or pitch movement) and whether naïve listeners can achieve this” (p.4). Given the possible inability of naïve raters and non-specialist teachers to distinguish separate features of prosody, along with disagreements in the literature about which features of prosody are influencing accentedness judgements, it was decided to treat prosody as one discreet unit in the questionnaire. While recognising that this decision may lead to the bundling of distinct features, such as rhythm and pitch, into one and therefore losing more detailed information about aspects of prosody which might influence accentedness ratings, it was felt this was necessary in order to focus on the overall goals of the study. Question 2b therefore posed the general sentence-level prosodic question as to what extent ‘sentence-level features’ (with the examples of ‘the ‘music’ of the sentences; the rhythm of the sentences’) sound native-like or not.

Along with phonological features of speech, there are extra-phonological features, which have also been identified as relevant to accentedness ratings. These are discussed below.

Question 2c: Fluency: Temporal Features

Saito et al. (2016) stress the importance of including what they describe as ‘temporal features’ in accentedness ratings, by which they mean features of fluency such as speech rate, pauses or hesitant speech. “A good speech rate,” they note, “should sound natural and be comfortable to listen to” (p.19). In a study by Davies (2019), certain features of the speech tokens were modified to assess what effect this had on listener feedback. In one block, for example, pitch contours were flattened to see if insufficient pitch contrast affected rater judgements of the tokens, while in another, speech rate was slowed. The findings suggested that “accentedness ratings were not systematically influenced by manipulations of pitch contours, but were influenced by speaking rate manipulations” (p.2). Similarly, Munro (1995) found that “differences in speaking rates, intonation patterns and timing” may have played a role in the listeners’ assessments” (p.17). Question 2c, therefore, takes this ‘fluency’ aspect of speech into consideration and asks about the temporal features of the tokens with the examples, “the speech is fluent; speaker is slow or hesitant”.

Question 2d: Vocabulary

Appel et al. (2019) in a study into the difference between comprehensibility and accentedness in L1 English raters’ evaluations of L2 English learners’ speech tokens concluded that, “a

substantial portion of variance in L1 English raters' assessments of comprehensibility and nativeness was attributed to lexical (and associated grammatical) features" (p.29). They studied three aspects of lexical richness, the first being, *lexical density*, which they described as the percentage of information-carrying 'content' words (like verbs and nouns) to 'function' words (such as determiners, auxiliaries and prepositions), which serve a primarily grammatical purpose. A larger percentage of content words, in this sense, indicates a higher lexical density. The other two aspects of lexical richness which they identified were *lexical variation* (the range of vocabulary used) and *sophistication* (the use of less common, more precise, or more specialized words based on frequency analysis from corpus data). They suggested that, in a picture description task, "comprehensibility tied exclusively to dimensions of lexical density and variation, whereas nativeness was more broadly associated with all three dimensions of lexical richness" (p.22). Trofimovich and Isaacs (2012) suggested that lexical errors, such as "semantically inappropriate words" and "imprecise vocabulary choice" (p.5), might also be influencing listener judgement, although they conclude that lexical errors might be more influential in comprehensibility ratings rather than in accentedness ratings. However, in view of this possible influence of aspects of lexis on listener judgements of accentedness, Question 2d asks about 'overall vocabulary features' with the examples, "odd use of words; good range of words".

Question 2e: Grammatical Accuracy

In Question 2a above, it was noted that Saito et al. (2016) argued that judgements of accentedness are primarily linked to pronunciation, particularly segmental features, rather than lexico-grammatical features. Trofimovich and Isaacs (ibid.) also commented that "grammatical errors and some aspects of vocabulary in L2 learners' speech appear to distract listeners from attending to the message, but have less relevance to their perception of L2 accent" (p.9). In contrast, this issue was discussed in Section 5.2.2.3, where it was proposed that it is the linguistic competence of speakers, as well as their accent, which influences NS judgements of comprehensibility and nativeness (Crowther et al., 2018; Piske et al., ibid.; Smith & Nelson, 1985). This was the reason given for choosing learners of B2 level and above as participants in this research, as it was felt their speech would be less likely to contain lexical and grammatical errors. However, in view of the possible influence of these errors in the speech tokens, it was decided to include Question 2e, which asks raters to judge 'grammatical accuracy' with the examples 'words in wrong order; no grammar mistakes'.

Question 2: Further Comments

At the end of Section 2 of the questionnaire, there was an optional section, inviting the rater to give further details regarding their judgement, including the examples: “sounds like an Italian accent, ‘th’ sound pronounced as ‘d’ ”.

Question 3: Native Speaker English Accent

In Section 5.2.1.1 above, it was discussed how the English variety of the listener (British English, Australian English, ...) might affect how the raters judge NNS speech tokens. This study is being conducted in an EFL context, that is to say, in a country where English is not the mother tongue. Portuguese speakers of English in Portugal are heavily influenced by American English through exposure to American films, series and so on, which are subtitled rather than dubbed. However, it is British English which largely dominates in education and the workplace. It was felt, therefore, that in this study, which works on British English, it was important to establish whether raters were making their nativeness judgements in relation to British English or another English variety for each token. Therefore, Question 3 asked raters, “which native speaker English accent does the speaker most closely approximate?” with the options: British English accent, American English accent, no native speaker accent, I don’t know and other, with space for comment.

4.2.3.3 Conclusion – Rating Speech Tokens

In a comprehensive study into variables affecting foreign-accented speech perception ratings, Kraut and Wulff (2013) concluded that “listeners’ perception of FAS [foreign-accented speech], rather than being solely determined by one factor alone, is shaped by a complex interplay of several characteristics of the FAS sample” (p.2). This is supported by Saito et al. (2017), who note that listeners attend not just to phonological features of speech, but to “a much broader range of linguistic dimensions (vowel/consonant errors, word stress, fluency, lexis, grammar) when judging comprehensibility” (p.14). Given this apparent complexity of variables affecting judgements of accentedness and the difficulty of teasing apart the concepts of comprehensibility and accentedness, questions on accentedness, along with more detailed questions on segmental/word-level, sentence-level, temporal, lexical, and grammatical variables, were included in the Accent Rating Questionnaire.

4.3 PARTICIPANTS

Participants for the study were recruited via an advertisement in the NOVA University, Lisbon Facebook page and circulated by the CLUNL (Linguistics Research Centre of NOVA University, Lisbon) and CETAPS (Centre for English, Translation and Anglo-Portuguese Studies) research groups. Prior to inclusion in the study, participants were asked to complete the Learner Consent Form and Pronunciation Questionnaire on Google Forms (Appendix D) to ensure that they fulfilled the research parameters.

Firstly, only applicants who had not been exposed to English before the age of 5 were chosen. This corresponds to the findings that place linear decline for pronunciation at 5-6 years of age or below (Bley-Vroman, 1989; Long, 1990; Birdsong, 2006) as discussed in Section 2.1 above. Three participants in Group 1 reported and five in Group 2 reported being first exposed to English between 5-6 years of age, but in all cases this was receptive exposure in the form of ‘watching English content with subtitles’ ‘watching content without subtitles’ and the more general ‘listening to spoken English’, with the exception of Participant 6, who mentioned ‘speaking in English’. Four participants in Group 1 and 2 participants in Group 2 reported first exposure to English between the ages of 7 and 11, and this exposure was more general including reading and speaking, with participant 1 citing ‘formal learning’, Participant 8 citing ‘listening to music’ and Participant 11 ‘playing videogames in English’. Only one participant, Participant 7, reported exposure starting between 10 and 17 years of age. Participant 7 reported European Portuguese (EP) as their L1, but was schooled in Venezuela. So, the majority of participants, five in Group 1 and five in Group 2 reported formal instruction starting between 7 and 11 years old.

Furthermore, as mentioned in Section 5.2.2.3, the chosen participants all reported a level of English of B2 or above according to the CEFR, in the hope that this would minimise or eliminate grammatical and lexical errors so that rater judgements would, therefore, be based on pronunciation features alone. The participants also confirmed that they would not be receiving other English language pronunciation intervention during the period of the course.

All participants indicated British English as their target, except for Participant 1 who initially chose American English, but said they were happy to work on British English. Participant 15, who initially indicated they were happy with their current accent, reported having lived in the UK and felt they could always ‘get better’. Participant 15 was therefore included in the study.

Finally, it was decided to include only speakers of L1 EP. This was partly to facilitate a comparative analysis approach between English and EP, which was adopted with Group 1 as described in Section 5.5.1 below. However, it was also hoped that this decision might avoid any possible listener bias

effects. In Chapter 3.1 above, the issue of accent prejudice was raised, noting that there is evidence to suggest some accents are seen as more desirable than others. There is some evidence of this also playing a role in accent evaluation. For example, Kang and Rubin (2009) reported on what they called, *Reverse Linguistic Stereotyping* (RLS), which appeared to show that a speaker's believed cultural background was influencing listener judgements. In their study, 158 participants were asked to evaluate the speech of three 4-minute segments of a lecture. The first and third segments were identical, delivered by the same NS English speaker, but in the first case, the raters were led to believe that Speaker 1 had a Euro-American background and appearance and Speaker 3 an East-Asian background and appearance. The second segment was presented as a distractor. The speaker of the third segment was found to be judged more harshly by listeners, with the authors commenting that listeners "heard" more accented speech when they were mistakenly made to believe they were listening to a native Chinese instructor" (p.452). In an attempt to neutralise possible RLS effects, it was decided to limit participants to L1 speakers of EP. Since all participants might therefore be expected to have a comparable L1 influence on their accent, possible RLS effects might be minimised, but no information regarding speaker profiles was passed on to the raters.

Volunteers who did not fulfil each of these criteria were dropped from the study prior to the intervention. Initially there were 16 participants in each group, giving a total of 32, but 5 dropped out prior to the start and 3 did not attend the course giving a total of 24 participants. During the course of the intervention, 5 participants were disqualified from the study as they did not attend a minimum of 8 of the 10 intervention sessions. A further 3 participants were not used in the analysis in order to ensure that Group 1 and Group 2 were the same size (8 participants), but the tokens of one of these participants (Speaker 0) were used as a practice exercise for raters before each rating session and results not included in the analysis. This left a total of 16 participants, 8 in each intervention group, whose profile is identified in Table 2 below. There were 4 male and 12 female participants with an age range from 18 to 51 and an average age of 28.

The Control Group was made up of B2 level students at Nova University, who were studying English Language as part of their course. They attended General English classes given by an NS of British English, but there was no pronunciation focus. They all reported EP as their L1, were aged between 18 and 38 when the first token was recorded and none reported exposure to English prior to the age of five. They provided speech tokens at the same time intervals as the participants attending the pronunciation intervention classes. This group consisted of 9 participants, but 4 did not complete all 3 tasks, so 3 more participants were added at a later stage and tokens collected under the same conditions, in order to ensure that the Control Group contained 8 participants to match the two intervention groups.

Table 2. Participant Profile

Group	Participant	Gender	Age at start of research	Nationality	Mother-tongue	AoA for L2 English	Self-reported English level
Group 1	1	F	49	Portuguese	European Portuguese	7-11	B2/C1
	2	F	28	Portuguese	European Portuguese	7-11	B2
	3	F	51	Portuguese	European Portuguese	5-6	C1
	4	M	21	Portuguese/ Costa Rican	European Portuguese	7-11	B2
	5	F	23	Portuguese	European Portuguese	5-6	C1
	6	F	18	Portuguese	European Portuguese	5-6	B2
	7	F	18	Portuguese	European Portuguese	10-17	B2
	8	F	41	Portuguese	European Portuguese	7-11	C1
Group 2	9	F	18	Portuguese	European Portuguese	5-6	B2
	10	F	28	Portuguese	European Portuguese	5-6	B2
	11	F	28	Portuguese	European Portuguese	7-11	C2
	12	M	27	Portuguese	European Portuguese	5-6	C2
	13	F	18	Portuguese	European Portuguese	5-6	B2
	14	F	18	Portuguese	European Portuguese	5-6	B2
	15	M	41	Portuguese	European Portuguese	7-11	unknown (lived in UK)
	16	F	20	Portuguese	European Portuguese	5-6	C1
Control	17	M	32	Portuguese	European Portuguese	7-11	B2
	18	F	19	Portuguese	European Portuguese	5-6	B2
	19	F	38	Portuguese	European Portuguese	7-11	B2
	20	F	20	Portuguese	European Portuguese	7-11	B2
	21	M	18	Portuguese	European Portuguese	5-6	B2
	22	M	23	Portuguese	European Portuguese	5-6	B2
	23	F	19	Portuguese	European Portuguese	5-6	B2
	24	M	18	Portuguese	European Portuguese	7-11	B2

The NS Group consisted of 3 NSs of British English resident in Portugal, 2 male and 1 female, aged between 41 and 59 at the start of the research, with tokens collected at the same time intervals as for the other 3 groups.

All participants were asked to sign a consent form and assured that all information collected during the study would be used for academic purposes only and would be treated anonymously (Appendix D). None of the participants received financial remuneration for their inclusion in the study, but all participants received a Certificate of Participation upon completion of the intervention.

4.4 INITIAL DATA COLLECTION AND RATING

Prior to the commencement of the courses, all participants were asked to fill out a Learner Consent Form and Pronunciation Questionnaire (Appendix D). The Learner Consent Form assured participants that that all contributions would be treated anonymously and asked them to indicate they agreed to the use of their data and voice recordings for scientific purposes (Question 1). They were asked for also for some personal information: their name (Question 2), their gender (Question 3), age (Question 4), email (Question 5), asked to indicate whether they preferred to attend on Friday or Saturday (Question 6), asked to commit to attending all the lessons (Question 7) and asked whether they would be attending other PI courses during the period of the course (Question 8). This was followed by a Pronunciation Questionnaire, which asked questions about their experience of English and their pronunciation goals. They were asked for their nationality (Question 1), their mother tongue (Question 2), their current level of English with proof (Question 3), the age at which they were first exposed to English (Question 4), what the initial exposure involved, such as watching TV or speaking English (Question 5), the age at which they started formal instruction (Question 6), a description of their current accent (Question 7). Question 8 asked them to identify if they had difficulty with any particular features of spoken English with the options: individual sounds (e.g. th sound; h at the beginning of a word), word stress (eg where to accent a longer word), sentence-level features (e.g. the music or rhythm of sentences) or fluency (slow or hesitant speech). They were allowed to tick as many boxes as they wished and were given space to add other features (Question 9). Finally, they were asked which accent they would like to have in English (Question 10). Questions 1,2,3,4,5, 6 and 10 were designed to ensure that participants complied with the participant profile, that is, native speaker of EP with B2+ level of English, who started learning English at the age of 5+ and whose stated pronunciation target was SBE, as discussed in Section 4.3. Questions 7,8 and 9 invited participants to describe their current accent in English so that it could be compared with the judgement of the raters (Section 4.2.3.2).

They were also asked to provide a speech sample in response to the following question:

Topic: Talk about a memorable holiday: what you did, who you were with, what you enjoyed/didn't enjoy and so on

These were then edited to approximately 30-second mp3 files as detailed in section 5.2 above. These tokens were then attached to the Accent Rating Questionnaire (Appendix C) and sent to the 9 native speaker raters.

4.5 INSTRUCTION PROCEDURE

During the intervention phase of the research (March – May 2022) participants were divided into 2 groups. The Group 1 intervention took place on Friday evenings and Group 2 on Saturday mornings, so participants were given the choice of which timetable suited them better. Those who had no preference were assigned to the group with fewer participants so that the size of the two groups was balanced. Materials for the intervention were produced by the author unless otherwise stated. Lessons were conducted via Zoom. The following gives a detailed overview of the two interventions.

4.5.1 Group 1 - Bottom-up Intervention

The first group received instruction aimed at raising learner self-awareness to their accent as proposed in Section 3.2.2, plus explicit pronunciation instruction. The instruction followed a FonFS approach (Section 3.3.1) based on a contrastive analysis approach (Section 3.3.3.4) focused on certain features of English pronunciation which have been identified as problematic for native speakers of EP. Choices regarding areas to address in this intervention were largely based on Shepherd (2001) and on my own experience of teaching English in Portugal over 30 years, but were also supported by other sources (Collins & Mees, 2009; Cruz-Ferreira, 1995, 1998, 2002; Rupp, 2018; Saritas & Saritas, 2022, Sedlackova, 2007). The intervention worked on both segmental and suprasegmental features (Section 3.3.2) as suggested by research in the area (Lee et al., 2015; Saito, 2012). The methodology attempted to include insights from research into perception and production training as identified in Section 4.2 and 4.3, with the materials largely based on the PPP approach used in *Teaching Pronunciation* (Celce-Murcia et al., 2010) as described in Section 3.3.3.1 above and included corrective feedback (Section 3.3.3.2), which research suggests promotes positive results (Gooch et al., 2016; Saito & Lister, 2012). After the first, lesson, each lesson started with a review of the content of previous lessons and gave learners the opportunity to reflect on the homework task. There follows a description of each lesson with the justification behind it and the resources used.

Lesson 1. Pronunciation Self-Awareness Raising Lesson and Suggestions for Target Accent Input

This was the first of two lessons (the other one being Lesson 10) which explicitly focused on encouraging the learners to listen to themselves. They were introduced to the notion that learners may not be good at assessing their own skills (Kruger and Dunning, 1999; Trofimovich et al., 2016). They were asked to introduce themselves to the group and were given peer and teacher feedback on their accent, which has been found to help raise learner self-awareness (Miller & Ng, 1994; Tsunemoto et al., 2022) (See Section 3.2.2). The quality of input has been suggested as an important factor in pronunciation attainment (Moyer, 2014a, Schmidt, 2012) (Section 2.3). A number of online resources were suggested which would enable learners to listen to their target accent – SBE, as follows: BBC ‘Pronunciation Workshop’ (https://www.bbc.co.uk/learningenglish/english/learn_pronunciation); listening to people with SBE accents - Benedict Cumberbatch; Keira Knightly; Daniel Craig ...; Pronunciation apps such as *Duolingo* and *Busuu* set to British English and British Council’s *Sounds Right*. Homework was to listen to their Pre-Intervention speech token to use as a diagnostic (Couper, 2015).

Lesson 2. L1 Interference – Possible Interference for L1 Speakers of EP in SBE: Consonants

Learners were introduced to the notion of L1 interference (Cruz-Ferreira, 1995; Shepherd, 2001). This lesson focused on two minimal pairs consonant contrasts as follows:

- voiceless dental fricative /θ/ and voiced dental fricative /ð/. EP has no dental fricatives and these have therefore been identified as problematic for L1 speakers of Portuguese (Shepherd, *ibid.*, Sedlackova, 2007; Rupp, 2018).
- voiceless palatal fricative /ʃ/ (*shoe*) and voiceless palatal affricate /tʃ/ (*chew*), identified as problematic as EP does not contain the latter (Cruz-Ferreira, *ibid.*; Shepherd, *ibid.*).

Instruction followed the PPP approach described in Section 3.3.3.1 which started with minimal pair perception training (<https://www.englishclub.com/pronunciation/minimal-pairs-t-th.php>), followed by controlled production training at the word-level and then at the sentence level. Materials were adapted from *English Pronunciation in Use* (Hancock, 2017, p.42). Learners were invited to practise the sounds at each stage and were given corrective feedback. The lesson focused on the /θ/ versus /ð/ distinction, but included suggestions for working on /ʃ/ and /tʃ/ (Hancock, 2003, p.158). Homework was to practise the two sets of minimal pairs with exercises introduced in the lesson.

Lesson 3. L1 Interference – Possible Interference for L1 Speakers of EP in SBE: Vowels

The lesson focused on minimal pair vowel contrasts as follows:

- Tense/lax vowel distinction with focus on the front, high, lax vowel /ɪ/ (*live*) and the front, high, tense vowel /i:/ (*leave*), identified as problematic for speakers of EP (Shepherd, *ibid.*) possibly because of a failure to perceive the two sounds as distinct (Rato et al., 2014) and a tendency to produce both as the Portuguese front, high vowel /i/ (Sedlackova, *ibid.*), as the tense/lax distinction does not feature in Portuguese (Rupp, *ibid.*).
- lax /ɒ/ (*shot*) and tense /ɔ:/ (*short*), which may be problematic for speakers of EP as they do not distinguish between the two sounds (Shepherd, *ibid.*).

Instruction followed the PPP approach described in Section 3.3.3.1 which started with minimal pair perception training, followed by controlled production training at the word-level and then, at the sentence level. Material used were from the English Club website (<https://www.englishclub.com/pronunciation/minimal-pairs-i-ee.php>) and Headway Pre-Intermediate Pronunciation (Bowler & Parminster, 2006 p.34). Learners were invited to practise the sounds at each stage and were given corrective feedback. The lesson focused on the /ɪ/ versus /i:/ distinction, but included suggestions for working on /ɒ/ and /ɔ:/. Homework was to practise the two sets of minimal pairs with exercises introduced in the lesson.

Lesson 4. Word/Sentence Stress and Rhythm:

Again, this lesson followed a PPP approach. It was stated that most Portuguese words stressed on second-to-last syllable (Cruz-Ferreira, *ibid.*), while English is a language with so-called ‘free-stress’ (Fudge, 2015, p.12), where word stress conventions are more erratic and therefore possibly problematic for speakers of Portuguese (Sedlackova, *ibid.*). Also, Portuguese speakers may give unstressed vowels (schwa /ə/) their full value (Shepherd, *ibid.*). The lesson looked at irregular stress patterns in English with the following example:

- Photograph /'fəʊtəgrɑ:f/; photographer /fə'tɒgrəfə/; photographic /,fəʊtə'græfɪk/ (adapted from Celce-Murcia, 2010, p.199)

Stressed syllables tend to be longer, louder and higher in pitch (Hancock, 2017, p.62)

Suggestions of resources for checking where primary and, if appropriate, secondary stress are placed with a word were given: Collins English Dictionary

(<https://www.collinsdictionary.com/dictionary/english>) and the Oxford Learner’s Dictionary

(<https://www.oxfordlearnersdictionaries.com/>). The lesson then looked at the use of schwa /ə/ in unstressed syllables using the following phrase:

- an apple and a banana /ən 'æpəl ænd ə bə'nɑ:nə/ (from English Pronunciation in Use, 2003, p.23)

Learners were invited to identify which syllables are stressed and which are unstressed in the phrase and this and similar phrases.

Moving on to sentence-level rhythm, it was discussed how English is considered a syllable-timed language (Collins & Mees, 2009), while EP shows features of both stress and syllable timing (Butler & Frota, 2018), which might lead to difficulties in this area (Collins & Mees, *ibid.*). The learners did a sentence-level rhythm drill exercise as follows:

- CATS CHASE MICE
The CATS have CHASED MICE
The CATS will CHASE the MICE
The CATS have been CHASing the MICE
The CATS could have been CHASing the MICE
From Celce-Murcia et al. (2010, p.210)

It was also noted that stress is used for emphasis in English where, sometimes, word order can be used in EP:

- I would prefer her to COME
Portuguese: Eu prefiro que ela venha
- I would prefer HER to come
Portuguese: Eu prefiro que venha ela

(From Cruz-Ferreira, 1998, p.173)

Homework was to use the Youglish website (<https://youglish.com/>) and to look for the phrase: *That's an interesting question.* They were asked to listen to different speakers of UK English and to choose one to copy.

Lesson 5. Intonation:

This lesson involved further work on stress-timing in English, specifically the fact that content words tend to be stressed, while function words are unstressed (Authentic Listening Resource Pack 1, Hancock, 2018). The lesson then moved on to intonation. Cruz-Ferreira, (2002)

suggest that intonation differences between SBE and EP may lead to communication breakdown (Cruz-Ferreira, *ibid.*). She states, for example, that the sentence *eles vivem em Lisboa* (p.13) (they live in Lisbon) which, with a rising tone at the end would be interpreted as a question in EP, but in English would be interpreted as a tentative statement since the question would be: *Do they live in Lisbon?*

Learners worked on nuclear stress, which highlights the feature of the sentence the speaker wants the listener to pay attention to (Gilbert, 2014) and is signalled by a change in pitch, increased volume and the lengthening of the vowel within the nuclear syllable, using the example of the question: *When do you begin your HOLiday?* This was then extended to a pairwork activity involving a conversation. Learners listened and commented on an example spoken by English NSs and then practised the dialogue in pairs:

A: When do you begin your HOLiday?

B: On the thirtieth of AUGust.

A: That's next SATurday!

B: We're leaving in the afterNOON.

A: And when are you coming BACK?

B: SATurday, September the thirteenth.

A: THIRtieth?

B: No, thirTEENTH!

(Celce-Murcia, 2010, p.200 from Hancock, 2003, p.52)

Learners then listened to the following passage from the Speech Accent Archive “*Please call Stella. / Ask her to bring / these things with her / from the store*” (http://accent.gmu.edu/browse_language.php?function=detail&speakerid=2362) and their attention was drawn to the way the speaker divided the text into thought groups (Celce-Murcia, 2010, p.221). Homework was to visit the Speech Accent Archive (http://accent.gmu.edu/browse_language.php?function=find&language=english) and listen to speakers 460 and 612 as examples of the use of rhythm and stress in SBE.

Lesson 6. Voiceless glottal fricative /h/ and Elision

This lesson focused on the tendency for Portuguese speakers to ‘drop’ /h/ in syllable initial position (Rupp, *ibid.*) or to overcompensate by adding /h/ prior to syllables with vowel in initial position (Shepherd, *ibid.*; Saritas & Saritas, 2022). As with the previous treatment of

segmentals, the lesson started with perception exercises and then moved on to an explanation of how to make the /h/ sound, followed by word- and then sentence-level production exercises (Celce-Murcia, 2010, p.81). The lesson then introduced weak /h/ and explained that it could be elided with weak function words: *he, him, his, her, have, had and has* (Rogerson-Revell, 2011, p.161). Homework was for learners to think about the feedback they had received so far and to use the exercises introduced to work on areas which they wished to modify.

Lesson 7. Orthographic Interference and Oral Reading for segmental practice

Learners were introduced to the notion of orthographic interference. Campos et al. (2018) suggest that Portuguese, while less transparent than Spanish in terms of grapheme-phoneme mapping, in English, the mapping is more opaque (See also, Shepherd, 2001). It was discussed how, in English, spelling was sometimes misleading, such as the pronunciation of ‘gh’: *tough* /tʌf/, *although* /ɔ:l’ðəʊ/, *ghost* /gəʊst/ and *hiccough* /’hɪkʌp/. Silent letters were also discussed such as in the pronunciation of the following: *climb* /klaɪm/, *knee* /ni:/, *castle* /’kɑ:səl/ and *autumn* /’ɔ:təm/. Then, the advantages and limitations of Oral Reading (Ricard, 1986) (Section 3.3.3.5) as a means of providing segmental training were discussed. The following text was given to practise Oral Reading with silent letters:

“It took me an hour to climb up to the castle. My muscles were sore and my knees hurt, but I knew it had been worth it. I saw the handsome castle in the evening light and felt calm as I listened to the autumn birds.” (Author’s own material)

Homework was to choose an oral reading passage to work on areas identified as non-target in feedback received so far.

Lesson 8. Coarticulation

A tendency for Portuguese speakers to omit features of sandhi variation and favour *citation forms*, the form in which they are pronounced when said as individual words, has been identified as a problem area for speakers of EP (Shepherd, *ibid.*) possibly due to orthographic interference (Section 2.2). Shepherd gives the example of consonant-consonants deletion such as *this stop* [ðɪstɒp] and *that time* [ðætəɪm], where L1 Portuguese speakers may fail to elide the repeated consonant in the phrase and pronounce them in citation form: *this stop* /ðɪs stɒp/ and *that time* /ðæt təɪm/. The lesson looked at various sorts of linking in English: *catenation* (consonant to vowel linking), for example *an egg* /ən ɛg/ is pronounced a/neg [ə’nɛg] in fluent

speech in SBE (Rupp, 2018), *intrusion* (the addition of an extra letter in fluent speech in SBE, for example *you ought* /ju: ɔ:t/ is pronounced *you (w)ought* [ju:wɔ:t] (Rogerson-Revell, *ibid.*, p.170) and *assimilation*, (when a sound in one word causes changes to sounds in neighbouring words), for example, in *public*/ in 'pʌblɪk/ become 'im public' [ɪm 'pʌblɪk] (Rogerson-Revell, *ibid.*, p.163). Learners were asked to add catenation and intrusion to the following sentence: *I used to think I'd be a teacher when I grew up, but as I got older, I began to ask myself if there was another option* (author's own materials) and practise reciting it. They also did an assimilation exercise to identify which sound changed in a series of sentences (Hancock, 2003, p.88). For example, in the phrase *I've got a bad cold*, the word *bad* sounds like *bag*. There was also a revision of elision, which was introduced in Lesson 6. Homework was to practise the materials looked at in class.

Lesson 9. General American (GA) vs. Standard British English (SBE) differences

As the target accent of the course was SBE and some learners were judged by the raters to have an American influence on their accent, it was deemed necessary to include this issue in the materials. There was a particular focus on the post-alveolar approximant /r/ in this regard. This has been identified as a significant area of difference between SBE and GA, which are nonrhotic and rhotic respectively (Celce-Murcia et al., *ibid.*). Learners listened to and practised the phoneme in both accents with a contrastive analysis exercise, involving words and phrases pronounced in both accents, for example *four days* /fɔ: deɪz/ in British English and is pronounced *four days* /fɔr deɪz/ in American English (from Hancock, 2007, p.52), with the added /r/ phoneme. Also addressed was the issue of *T-voicing* in GA, when /t/ mutates to a voiced flap under certain conditions, leading to the /t/-/d/ contrast being neutralised (Boberg, 2015) and the SBE back low rounded vowel /ɒ/ versus the more central and unrounded GA /ɑ/ vowel, heard in the production of *stop watch* as /stɒp wɒtʃ/ (SBE) versus /stap wɑtʃ/ (GA) (Celce-Murcia et al., *ibid.*). As before, these differences were presented first in terms of perception and then moved on to controlled and less controlled exercises. Homework was to practise the following test with an SBE accent:

“While there may be certain differences between British and American English, the key takeaway is that the two have more similarities. Accidentally using one instead of the other will not automatically lead to problems.”

Adapted from: <https://www.britishcouncilfoundation.id/en/english/articles/british-and-american-english>

Lesson 10. Review of course content.

This lesson started with a review of all the topics covered in the course. Then the 30-second section of token 1 which had been sent to raters was played for each learner, who were invited to comment on their pronunciation (Couper, 2015) and then received peer, teacher and rater feedback (Miller & Ng, 1994; Tsunemoto et al., 2022), both of which have been proposed as being helpful in raising learner phonological self-awareness (Section 3.2.2). This was designed to give learners feedback on their pronunciation in order to complement Lesson 1 and work on pronunciation self-awareness.

4.5.2 Group 2 - Top-down Intervention

The second group received instruction aimed at raising learner self-awareness to their accent, which was identical to that delivered to Group 1 in Lessons 1 and 10. However, the rest of the intervention focused on giving the learners an awareness of the influence of the social and communicative context on their accent (LaScotte et al. 2021; Pennington, 2019) as discussed in Section 3.4 and offer them strategies which might help them work on their L2 pronunciation beyond the level of segmental and suprasegmental features to think more holistically of their production in general, as proposed in Sections 3.4.2, 3.4.3 and 3.4.4. As there are a limited number of teaching resources which target this area, resources used were designed by the author and supplemented with online materials as indicated in the lesson descriptions. It is important to state that there was a certain degree of overlap between the two groups, such as in the inclusion of the influence of L1 interference on L2 pronunciation and the use of oral reading as a strategy but, with Group 2, the focus was on overall performance, rather than structural ‘accuracy’. Corrective feedback (Section 3.3.3.2), which research suggests promotes positive results (Gooch et al., 2016; Saito & Lister, 2012) was provided on each task. After the first, lesson, each lesson started with a review of the content of previous lessons and gave learners the opportunity to reflect on the homework task. There follows a description of each lesson with the justification behind it and the resources used.

Lesson 1. Pronunciation Self-Awareness Raising Lesson and Suggestions for Target Accent Input

This was the first of two lessons (the other one being Lesson 10) which explicitly focused on encouraging the learners to listen to themselves and was a replica of the lesson delivered to Group 1. They were introduced to the notion that learners may not be good at assessing their own skills (Kruger and Dunning, 1999; Trofimovich et al., 2016). They were asked to introduce

themselves to the group and were given peer and teacher feedback on their accent, which has been found to help raise learner self-awareness (Miller & Ng, 1994; Tsunemoto et al. (2022). (Miller & Ng, 1994; Tsunemoto et al. (2022) (See Section 3.2.2). The quality of input has been suggested as an important factor in pronunciation attainment (Moyer, 2014a, Schmidt, 2012) (Section 2.3). A number of online resources were suggested which would enable learners to listen to their target accent – SBE, as follows: BBC ‘Pronunciation Workshop’ (https://www.bbc.co.uk/learningenglish/english/learn_pronunciation); listening to people with SBE accents - Benedict Cumberbatch; Keira Knightly; Daniel Craig ...; Pronunciation apps such as *Duolingo* and *Busuu* set to British English and British Council’s *Sounds Right*. Homework was to listen to their Pre-Intervention speech token to use as a diagnostic (Couper, 2015).

Lesson 2: Articulatory Settings in English

This lesson introduced the concept of articulatory settings (Section 3.4.4). Learners listened to a University of Albeda voice specialist, David Ley, talk about features of SBE speech (<https://youtu.be/ldAKIzq7bvs>). This was supported by the presentation of a list important features of SBE voice quality:

- *general lack of muscularity (lax)*
- *neutral character of lip movement*
- *preference for centralised vowels (/ə/)*
- *many articulation with tongue tip on alveolar ridge (t/d/l ...)*
- *return to rest position within articulations*

From Honikman (1964) in Rogerson-Revell (2011, p.38)

Learners were invited to think of *cognates*, words which are similar in English and Portuguese such as: *banana*, *cinema*, *natural*, *crime* and *familiar* and to listen to them pronounced in English and Portuguese, to produce word pairs themselves and to comment on how the voice quality was different (Thornbury, 1993). This was followed with other ideas for working on Articulatory Settings:

- *Write down a few common phrases in Portuguese and translate them to English*
- *Articulate one of the phrases without making any sound. Can we guess if it is English or Portuguese?*

- Say one phrase pair out loud. Can listeners explain differences in voice quality, e.g. volume, nasal quality, length of sounds, pitch (high/low), other features?
 - Say one of your Portuguese phrases with an English accent
 - Try an English phrase with an exaggerated Portuguese accent
- (Adapted from Thornbury, 1993; Rogerson-Revell, ibid.)

Homework was to reflect on whether they needed to modify their voice quality in English.

Lesson 3: Top-Down Approach to Pronunciation

Learners were introduced to the idea of a top-down approach to pronunciation with a quote from Pennington (2019b): “Because speech occurs in a connected stream of sound, it is more realistic to teach pronunciation “top-down” ... in context, rather than teaching only individual sounds in isolation” (p.1). Learners were then asked to listen to part of an interview with British actress, Emma Watson (<https://youtu.be/PBCOgyPv-nM>) and try to copy some phrases: “*It doesn’t LOOK like that*”, “*I thought I looked aMAzing*”, “*I made my parents CRAzy*”. This activity was to practise the Echo Method (Chung, 2017) described in Section 3.4.2 and lead to a discussion of stress, rhythm and intonation in her speech. The lesson went on to suggest other strategies for working on prosody, such as listening to songs, with the example of ‘*Lions and Tigers and Bears, Oh My!*’ from the Wizard of Oz (suggested by Gilbert, 2001) and using a limerick:

*There was an old man from Peru
Who dreamed he was eating his shoe
He awoke in the night
In a terrible fright
And found it was perfectly true!*
(In Celce-Murcia et al., 2010, p.351):

Homework was to find an interview with a British celebrity with SBE, such as Benedict Cumberbatch, Keira Knightly, Daniel Craig, Stephen Fry or J K Rowling and to try to copy short passages of their speech.

Lesson 4: Pronunciation as Skill Acquisition

Learners were introduced to the concept of pronunciation as a skill and the importance of developing new muscular habits and practising correctly (Fraser, 2006; Rogerson-Revell, *ibid.*) discussed in Section 3.1. In order to work on this, learners were introduced to the Bowen technique (Bowen, 1975) for working on areas which they felt needed modification. The example given was the /ɪ/-/i:/ distinction using the exercise below:

COLUMN A

1. Shall I heat the potatoes?

Shall I hit the potatoes?

2. Can you feel the bottle?

Can you fill the bottle?

3. He bit the banana

He beat the banana

4. Can you fetch the bins, please

Can you fetch the beans, please

5. Did you slip on the floor?

Did you sleep on the floor?

COLUMN B

a) Yes, I don't like them cold

b) Yes. Are you making 'batatas a murro'?

a) Oh yes! It's really cold!

b) Yes. I'll put more water in it

a) Did it taste nice?

b) Was he making a milkshake?

a) Yes, I'll get them now

b) Yes, are you cooking some for dinner?

a) Yes, someone must have just cleaned it

b) Yes, the bed was uncomfortable

Based on the Bowen Technique (Bowen, 1975) in Celce-Murcia (2010, p.148)

The teacher read from Column A which were shadowed by students. Then Student A read one sentence from Column A and Student B would attempt to give the appropriate reply from Column B. The exercise was designed to practise shadowing (Section 3.4.2) within a FonF paradigm (Section 3.3.1) with peer and teacher feedback. Homework was to use Youglish (<https://youglish.com/>) to practise features of production which learners wanted to work on, as identified in Lesson 1, and emphasising the need for repetition in order to work on skill acquisition (Section 3.1).

Lesson 5: L1 Interference

Learners were introduced to the notion of L1 interference from EP (Cruz-Ferreira, 1995; Shepherd, 2001) with examples given of possible problem areas: consonants: th (/ð/ this - /θ/ think); ch/sh (/tʃ/ chip - /ʃ/ ship); /h/ at syllable start (happy; hotel); vowels: tense/lax distinction /i:/ (sheep) - /ɪ/ (ship); /ɔ:/ (warned) - /ɒ/ (wand); half vowels: /r/ (brother; rest); prosody: non-use of schwa /ə/.

They listened to the following passage from the Speech Accent Archive, spoken by a speaker with L1 EP and asked to identify what features of possible L1 interference :

“Please call Stella. Ask her to bring these things with her from the store: Six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.”

36-year-old male Portuguese speaker from The Speech Accent Archive

(http://accent.gmu.edu/browse_language.php?function=detail&speakerid=513)

Learners were then asked to listen to the same text, but spoken by a 48-year-old British male (http://accent.gmu.edu/browse_language.php?function=detail&speakerid=2362). Learners were sent to breakout rooms in pairs to copy the first two sentences and then the groups were brought back together for feedback. This task was designed to practise the *golden sentence* strategy described in 3.4.2. For homework, learners were encouraged to find a speaker whose accent they liked in the Speech Accent Archive, and to copy the first two sentences.

Lesson 6: Language Aptitude and Social Context

This lesson attempted to raise learner awareness to personal traits, which might influence pronunciation outcomes and to think about social context, cited by Pennington (2019a) as fundamental to the top-down approach (Section 3.4). Learners were asked to speculate what characteristics made some people better second language learners than others. Issues of aptitude such as musical and mimicry ability (Section 2.4) were discussed, it was suggested that “... improvement is arguably within anyone’s reach, and ‘talent’ may be less influential than one’s overall orientation” (Moyer, 2015, p.10).

It was then discussed how certain personality traits, such as extraversion and empathy might promote pronunciation modification, while others might hinder it, such as language anxiety (Section 2.6). They were told about the research of Guiora et al. (1972) which suggested that speakers produced more accurate Thai sound when they were given alcohol (Section 2.6), so pronunciation modification might involve a degree of ‘letting go’ or “being someone else” (Pennington, 2019, p.377) as discussed in Section 3.4.3. The lesson finished with a questionnaire as follows:

- 1) *Are you a good mimic? Can it be taught? We will look at ways to train mimicry.*
- 2) *Are you motivated to sound like a native speaker? Why is this important to you?*
- 3) *Do you need to 'let go' of your L1 identity? Is it holding you back? We will look at ways to overcome this.*
- 4) *Do you lack 'language self-confidence'? Is this affecting what you believe you can achieve?*

The purpose of this lesson was to address learner motivation (Section 2.7.1) in the sense of assuring them that, given the right attitude, progress can be made in accent modification and to raise personality and identity issues which might be affecting their pronunciation (Sections 2.6 and 2.7). Homework was to reflect on issues brought up in the lesson.

Lesson 7: Shadowing/Oral Reading (PPR)

Learners were introduced to *shadowing* (Section 3.4.2) - a mimicry activity where learners imitate a presented speech stimulus “as closely and quickly as possible” (Luo et al. 2008, p. 4) though the repetition can be near simultaneous or have a small delay. Alice in Wonderland from the BBC Learning English website (<https://www.bbc.co.uk/learningenglish/features/rama/alice-ep01>) was given as the text:

Narrator: *“Hello. This is the story of an ordinary girl called Alice. Our story begins on an ordinary summer's day. Let's meet Alice: here she is, sitting next to the river with her sister on a very hot summer's afternoon. Her sister is reading a book, and with no-one to talk to, Alice is bored... and a bit sleepy. Soon, she'll be in a Wonderland full of strange and wonderful adventures... but Alice doesn't know that yet...”*

Learners were paired in breakout rooms on Zoom and invited to take turns shadowing the text and giving each other feedback. Other resources were suggested for shadow listening from BBC Sounds podcasts (<https://www.bbc.co.uk/sounds/podcasts>). It was also suggested that shadowing could be practiced with songs, with the example of *An Englishman in New York* by Sting. Learners were provided with the lyrics and were invited to shadow the song:

*I don't drink coffee, I take tea, my dear
I like my toast done on one side
And you can hear it in my accent when I talk
I'm an Englishman in New York*

Learners were also introduced to peer-reviewed pronunciation reading (PPR) (Luo, 2016) (Section 3.4.2) where learners recite a text focusing on pronunciation. It was suggested that learners could ask someone with a good knowledge of English to listen to them, since NNSs have been shown to have good accent rating skills (Section 4.2.1.1) or they could record themselves and evaluate their own performance as suggested in Lesson 1. The British Council Audio Zone (<https://learnenglish.britishcouncil.org/general-english/audio-zone>) and British Council Podcasts for Professionals (<https://learnenglish.britishcouncil.org/business-english/podcasts-for-professionals>) were suggested as resources where podcasts were available with a transcript, so the resource could be used for both shadowing and PPR. Homework was to look at the resources suggested in the lesson and to practise one of the strategies suggested.

Lesson 8: Accent and Identity

This lesson aimed to give learners an awareness of identity as an important influence on accent (Section 2.7). It started with a quiz to encourage them to think about their accent in both their L1 and L2 English:

1. *In Portuguese, why do you have the accent which you have? Does it come from your parents? Your friends? Somewhere else?*
2. *Do you use the same accent (in Portuguese) when talking to your parents, your friends, your boss, the cleaning lady, ... ?*
3. *Do you change your accent when speaking to foreigners in Portuguese?*
4. *In English, why do you have the accent which you have? Where does it come from?*
5. *Would you like to be mistaken for a native speaker of English?*

Learners were introduced to the following quote: “Changing pronunciation means altering psychological, social and cultural aspects of the self ... a matter of openness and willingness to change” (Pennington, 2019, p.375) and told about mirroring (Myers, 2016) discussed in Section 3.4.3. Mirroring was described as follows:

1. *Choose a passage of speech by your target speaker e.g. a Ted Talk*
2. *Listen to the segment several times, noting important features of the speech e.g. stressed sections, pauses*
3. *Practise copying your model.*
4. *Perform it for someone/video it! Ask your listener(s) for feedback/decide what you want to change*
5. *Practise again ...*

Learners then watched an interview with Prince William from the English Speeches website (<https://englishspeecheschannel.com/english-speeches/british-royal-family-speech/>) and worked in pairs to mirror the opening sentences:

INTERVIEWER: *So, a real insight into why you and Prince Harry set up the Royal Foundation. Did you ever imagine it (would) grow like this?*

WILLIAM: *I think when we first started it, it started as a very small idea with Harry and I scratching our heads going ... how, er how can we do something that's going to help us in the future*

They were also introduced to the concept of the *golden speaker* (Ding et al., 2019) (Section 3.4.2) which suggests that they should choose a target from a similar background and whose voice is similar to their own. Homework was to reflect on their identity as a speaker of English and to choose a target to mirror.

Lesson 9: Using a Template/Channeling

In this lesson, learners were introduced to the idea of a 'template' (Gilbert, 2014), discussed in Section 3.4.2, which involves perfecting one sentence in the target accent. Learners listened to the phrase: *Good morning! How are you today?* on Youglish (<https://youglish.com/>), which allowed them to hear various renditions of the target phrase. They were then invited to choose one version, to listen to it several time and then to try it themselves. As in previous exercises, they were told to record themselves and/or ask for listener feedback until they were happy with the result. Then, learners were told about *channeling* (Moreno, 2016), the idea of pretending to be someone else in a roleplay situation, discussed in Section 3.4.3. They were given the following instructions and asked in practise in breakouts rooms on Zoom:

You meet a British colleague.

- 1. Invent a short conversation starting "Good morning! How are you today?"*
- 2. Act out the interaction.*
- 3. Then swap roles and try again. The British people should 'channel' their best British accent!*
- 4. Give each other feedback on your British accents.*

Each pair was then invited to perform their conversation for the class.

Lesson 10: Review of course content and Feedback

This lesson started with a review of all the topics covered in the course. Then the 30-second section of token 1 which had been sent to raters was played for each learner, who were invited to comment on their pronunciation (Couper, 2015) and then received peer, teacher and rater feedback (Miller & Ng, 1994; Tsunemoto et al., 2022), both of which have been proposed as being helpful in raising learner phonological self-awareness (Section 3.2.2). This was designed to give learners feedback on their pronunciation in order to complement Lesson 1 and work on pronunciation self-awareness. This was a repetition of Lesson 10 delivered to Group 1.

4.5.3 Control Group

The Control Group attended English language classes delivered by an English NS for the 10 weeks of the intervention, but received no explicit pronunciation instruction.

4.6 POST- AND DELAYED POST-INTERVENTION COLLECTION AND RATING OF SPEECH TOKENS

In early June 2022, on completion of the intervention, the participants were once again asked to send a speech sample, this time in response to the following question:

Talk about your favourite free time activity (hobby): what does/did it involve, what do you like about it, how did you start doing it, who do/did you do it with and where, what is your best memory of this activity and so on

These were then sent to the raters at the end of June 2022. Participants were also asked to complete the End of Course Feedback form (Appendix G), which asked whether they felt the course had increased their awareness of their own accent (Question 1), which features they had an increased awareness of (Question 2), which accent they would like to have in English (Question 3), if they felt the course had helped them to change their pronunciation (Question 4), which aspects of their pronunciation it helped them change (Question 5) and which activities they found most useful (Question 6).

A final collection of speech samples was completed in September 2022. The prompt was:

Talk about an animal/animals: do you have/have you ever had any pets, what are/were they like, any memorable animal stories, your favourite/least favourite animal, which animal would you like to be and so on.

The last set of speech samples was sent to the raters at the end of September 2022 and ratings collected.

As previously mentioned, due to insufficient tokens from the Control Group, a further period of speech sample collection was conducted, in April 2023, June 2023 and September 2023 with B2 level learners from Nova University as before. The raters remained the same throughout.

This chapter has provided an explanation of research questions and a detailed justification of decisions made regarding methodology, namely the choice of raters, tokens, rating methods and choice of participants. There followed a precise description of the intervention delivered to the two intervention groups. The following chapter looks at the findings of the study as they relate to the research questions.

5. RESULTS

“L2 pronunciation teaching and materials development are full of talented craftspeople who bring divergent, yet under-researched approaches to teaching, and who may show little knowledge of research findings. On the research side, research articles are often written with pedagogical implications that no one applies.”

(Levis, 2017, p.1)

This chapter gives a detailed description of the results obtained for each research question in turn, along with a discussion of the results and conclusions to be drawn in each case. This is followed by a general discussion of the results, including possible limitations of the study and suggestions for future research.

5.1 QUESTION 1: TO WHAT EXTENT DO LEARNERS' PRE-INTERVENTION ASSESSMENTS OF THEIR OWN SPONTANEOUS SPEECH PRODUCTION CORRESPOND TO THOSE OF NS RATERS?

To answer this question, spontaneous speech tokens were collected from the 16 participants in the intervention groups (Group 1 and Group 2) prior to commencement of the intervention (Section 4.4 above). At the same time, participants were asked to complete the Learner Consent Form and Pronunciation Questionnaire (Appendix D) (Section 4.4), part of which asked them to describe their accent (Appendix D - Question 16 ‘how would you describe your current accent in English’) and to note if they had difficulty with individual sounds, word stress, sentence-level features and/or fluency in spoken English (Appendix D – Question 17). They were also given the option to add additional information (Appendix D – Question 18 ‘Add more details about the above question if necessary’).

NS raters were asked to rate the speaker tokens for nativeness on a 1-9 Likert scale from 1 (‘not native speaker’) to 9 (‘native speaker’) (Appendix C – Question 2), and to indicate whether the accent was influenced by any NS accent in particular (Appendix C - Question 3 ‘Which native speaker English accent does the speaker most closely approximate?’). They were then asked to rate specific features of the token: ‘pronunciation of individual sounds/words’ (Appendix C – Question 2a), ‘sentence-level features’ (Appendix C – Question 2b), ‘overall fluency’ (Appendix C – Question 2c), ‘overall vocabulary features’ (Appendix C – Question 2d) and ‘overall language accuracy’ (Appendix C – Question 2e) using the Likert scale as described above. Raters were also given the option to add additional information at the end of Question 2 (Appendix C – Question 2 ‘Give further details about your decisions above’).

The results of the participant and rater feedback were then compared, with participants' description of their current accent in English compared to the raters' nativeness judgement and the influence on their accent. Then, the participant and rater feedback for the three specific features of pronunciation: individual sounds/words, sentence-level features and overall fluency were compared. It was predicted that, in line with the predictions of the Dunning-Kruger effect (Kruger & Dunning, 1999), discussed in Section 2.4.1.4, many learners would judge their accents more favourably than would NS raters.

5.1.1 Speaker and Rater Judgements Compared

Table 3. Speaker 1 – Pre-Intervention comparison between speaker and rater accent judgements

Speaker 1	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NNS accent	4/9
Influence on Accent	Portuguese with American influence	5 raters - NNS accent 2 raters – I don't know 1 rater – American 1 rater – Iberian
Individual sounds/word stress difficulties	Yes	4/9 2 raters = /h/ omission in 'holiday' 1 rater = 'th' ('think pronounced sink')
Sentence-level difficulties	No	5/9
Fluency difficulties	No	5/9 1 rater = 'hesitant'

Speaker 1 identified their accent as a 'Portuguese accent mixed with some American accent'. Only one rater identified what they described as an 'Iberian' accent, but five raters identified a NNS accent and one rater identified an American influence on the accent. Ratets identified sentence-level and fluency difficulties not mentioned by speaker, with one rater describing fluency as 'hesitant'. However, in terms of nativeness and individual sounds/words there was agreement between rater and speaker that the speaker had problems in this area. Two raters mentioned omission of /h/ in the word 'holiday' /'hɒlədeɪ/ and one rater stated that 'think' /θɪŋk/ was pronounced as 'sink' [sɪŋk], but these were not specifically mentioned by the speaker. Overall, therefore, there is partial agreement between speaker and raters for Speaker 1.

Table 4. Speaker 2 – Pre-Intervention comparison between speaker and rater accent judgements

Speaker 2	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NNS accent in English	2/7
Influence on Accent	Portuguese	5 raters - NNS accent 2 raters – American 1 rater – Iberian 1 rater – I don't know
Individual sounds/word stress difficulties	Yes	3/7 2 raters = [th] ('th as d'/'de instead of the') 1 rater = /tʃ/-/ʃ/ distinction ('sh instead of ch in chocolates')
Sentence-level difficulties	Yes	4/7
Fluency difficulties	Yes	4/7 1 rater = hesitant

Speaker 2 described their accent as a 'Portuguese accent in English'. One rater identified an 'Iberian' accent, while five raters described the accent as NNS indicating agreement between speaker and raters. The speaker also identified problems with individual sounds/words but did not specify which ones. Raters agreed that this was a problem with a median rating of two on the Likert scale and two raters identifying 'th'/'ð/' as being replaced by 'd'/'d' in 'the' / ði:/, while one rater specified the /tʃ/-/ʃ/ distinction saying that 'chocolates' /'tʃɒkləts/ was pronounced [ʃɒkləts]. Sentence-level and fluency difficulties were also identified by both speaker and raters meaning that there was agreement between the two regarding this accent.

Table 5. Speaker 3 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 3	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	I don't know what I sound like	3/9
Influence on Accent	I don't know what I sound like	4 raters = NNS accent 2 raters = American 1 rater = Brazilian 1 rater = Iberian + British 1 rater = British
Individual sounds/word stress difficulties	No	3/9 4 raters = /i:/-/ɪ/ distinction

		(‘pissful not peaceful’; ‘short i in peaceful’)
Sentence-level difficulties	Some sequences of words	4/9
Fluency difficulties	No	3/9 1 rater = ‘hesitancy’

Speaker 3 stated ‘I don’t know what I sound like’. They received a median rating of three on the Likert scale from the raters for nativeness, meaning this was identified as an NNS accent, with American, Brazilian, Iberian and British offered as possible influences on accent. The speaker did not identify difficulty with individual sounds/word stress, but the raters gave a median score of three on the Likert scale, with four raters specifically commenting on problems with the /i:/-/ɪ/ distinction, with three stating that ‘peaceful’ /‘pi:sfəl/ was pronounced as ‘pissful’ [‘pɪsfəl] and one commenting ‘short i in peaceful’. Raters identified fluency difficulties, which were not seen as a problem by the speaker, but there was speaker/rater agreement that the speaker showed sentence-level difficulties. On the whole, however, there was little agreement between speaker and raters, with the speaker stating that they did not know what they sounded like.

Table 6. Speaker 4 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 4	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NS accent	6/9
Influence on Accent	American English	4 raters = American 3 raters = NNS accent 1 rater = British + Iberian 1 rater =Scandinavian or Polish
Individual sounds/ word stress difficulties	Yes	6/9
Sentence-level difficulties	No	7/9
Fluency difficulties	Yes	7/9

Speaker 4 described their current accent as an ‘American English accent’ and four raters agreed that there was an American influence on the accent, but there was some inter-rater disagreement with three raters describing this as merely an NNS accent, one rater identifying a British and Iberian influence and one rater describing it as a Scandinavian or Polish accent. Both speaker and raters identified individual sounds/word stress difficulties as a problem area, with raters giving a median Likert score of 6. Raters gave the speaker a high median score of seven for sentence-level features and fluency, but the speaker

indicated that they thought they had difficulties with fluency. Overall, there was close agreement between speaker and raters regarding this accent.

Table 7. Speaker 5 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 5	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NS accent	4/9
Influence on Accent	American English	5 raters = NNS accent 2 raters = American 1 rater = Spanish 1 rater = I don't know
Individual sounds/ word stress difficulties	No	5/9 2 raters = /h/ omission ('Hs dropped'; 'an ostel instead of a hostel')
Sentence-level difficulties	No	5/9 1 rater = intonation
Fluency difficulties	Yes	5/9 1 rater = 'hesitant'

Speaker 5 identified their accent as an 'American English' accent, but six raters judged it as an NNS accent with one specifying a Spanish accent and there was a median rater score of four for nativeness. Raters also identified problems at the individual sounds/word stress level with a median rating of five on the Likert scale for both features. Elision of /h/ at the start of a word was mentioned by two raters with one rater stating that 'a hostel' /ə 'hɒstəl/ was pronounced [ən 'ɒstəl]. The speaker did not identify sentence-level difficulties, but this was given as a median score of five by the raters with 'intonation' specifically mentioned as a difficulty by one rater. The speaker identified fluency as a problem area and the raters agreed with a median rating of five and one rater specifically mentioning hesitancy. However, overall, there was little agreement between speaker and raters.

Table 8. Speaker 6 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 6	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	I don't know what I sound like	5/9
Influence on Accent	I don't know what I sound like	7 raters = NNS accent 2 raters = American
Individual sounds/word stress	No	6/9

difficulties		
Sentence-level difficulties	Yes	5/9 1 rater = 'rise and fall of sentences'
Fluency difficulties	No	6/9 1 rater = 'a bit slow'

Speaker 6 stated: 'I don't know what I sound like'. Raters gave this accent a median score of five for nativeness with seven describing it an NNS accent with two raters identifying an American influence. Despite being unaware of their current accent, the speaker identified sentence-level difficulties, which was in agreement with the raters who gave a median rating of five on the Likert scale and one rater highlighting the 'rise and fall of sentences' as a feature. They did not elaborate, but this suggests that they identified NNS intonation in this speaker. The raters also highlighted difficulties at the level of individual sounds/word stress and fluency, giving both a median score of six on the Likert scale and one rater describing fluency as 'a bit slow'. This speaker was therefore classed as an NNS in all areas by raters, but the speaker stated they did not know what they sounded like.

Table 9. Speaker 7 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 7	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	I don't know what I sound like	4/9
Influence on Accent	I don't know what I sound like	3 raters = American 3 raters = NNS accent 1 rater = Brazilian 1 rater = Spanish 1 rater = I don't know
Individual sounds/word stress difficulties	Yes	5/9 2 raters = 'favourite' ('I found it hard to understand 'favorite'; 'favourite') 1 rater = 'consonants dropped'
Sentence-level difficulties	Yes	6/9 1 rater = intonation
Fluency difficulties	No	6/9

Speaker 7 stated ‘I don’t know what I sound like’. Raters were not in agreement regarding this accent. It received a median rating of four for nativeness, but it was identified as American-influenced by three raters, NNS by three raters, Brazilian and Spanish by one rater each and the final rater was unable to identify the accent. The speaker identified individual sounds/words stress and sentence-level difficulties and raters agreed, giving a median score of five and six respectively. The word ‘favourite’ is identified by two raters with one just writing the word ‘favourite’ in the comments and another stating ‘I found it hard to understand ‘favorite’’, suggesting comprehensibility problems. Neither rater specified the problem, but in the token, the speaker appears to pronounce ‘favourite’ /'feɪvərɪt/ as [ˈfɜ:vərɪt]. One rater also commented on ‘consonants dropped’ but did not elaborate. This probably referred the phrase ‘it was only three days’, where ‘it was only’ /ɪt wɒz 'əʊnli/ is pronounced [ɪ: wɒz 'əʊnli] with the /t/ of /ɪt/ dropped. This is a feature of rapid NS speech, but the change from lax /ɪ/ changed to tense /i:/ is not and this appears to have been the reason for the rater comment. Fluency was not identified as a difficulty by the speaker, but received a median of six from raters, suggesting it was not at NS level. The raters were not in agreement regarding the influence on the accent but agreed with the speaker regarding individual sounds/words stress and sentence-level difficulties. However, the speaker stated they did not know what they sounded like.

Table 10. Speaker 8 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 8	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NNS accent in English	3/9
Influence on Accent	Portuguese	6 raters = NNS accent 1 rater = Iberian 1 rater = British 1 rater = American
Individual sounds/word stress difficulties	Yes	4/9
Sentence-level difficulties	No	4/9
Fluency difficulties	No	3/9 2 raters = ‘hesitant’

Speaker 8 described their accent as a ‘Portuguese accent in English’ and while raters agreed that this was an NNS accent giving a median score of three for nativeness and seven raters describing it as an NNS accent, they did not specifically identify it as Portuguese, although one rater described it as ‘Iberian’. Raters judged this speaker quite harshly for the individual measures, with two raters specifically mentioning hesitancy for fluency. The speaker, however, only identified one measure – individual sounds/word stress – as problematic. This means that while speaker and raters judged this as

an NNS accent, they differed in their rating of sentence-level and fluency difficulties meaning there was partial agreement between speaker and raters.

Table 11. Speaker 9 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 9	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	I don't know what I sound like	4/9
NS Influence on Accent	I don't know what I sound like	5 raters = NNS accent 2 raters = American 1 rater = British 1 rater = Iberian
Individual sounds/word stress difficulties	Yes	4/9
Sentence-level difficulties	No	4/9 1 rater = 'words easy to understand but meaning less so'
Fluency difficulties	Yes	5/9 1 rater = 'hesitant'

Speaker 9 stated 'I don't know what I sound like'. Six raters described Speaker 9 as having an NNS accent, with one rater specifying 'Iberian'. Two raters identified an American influence, while one identified a British influence. The speaker noted individual sounds/word stress and fluency difficulties which were also noted by the raters who gave a median score of four and five respectively with one rater describing speech as hesitant. The speaker did not identify sentence-level difficulties, but this was identified as a problem by the raters who gave a median score of four, with one rater specifying 'words easy to understand but meaning less so' suggesting intelligibility difficulties beyond the level of pronunciation, but the rater gave no further details. The speaker, therefore, did identify two measures where they experienced difficulties in agreement with raters, but overall, they stated that they did not know what they sounded like.

Table 12. Speaker 10 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 10	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NNS accent	3/9
Influence on Accent	Portuguese	8 raters = NNS accent 1 rater = Iberian

Individual sounds/word stress difficulties	Yes	3/9 1 rater = ‘mammal sounds like memo’
Sentence-level difficulties	Yes	4/9
Fluency difficulties	No	4/9

Speaker 10 described their accent as a ‘Portuguese accent in English’, which agreed with the raters who gave a median score of three for nativeness, with all raters describing this as an NNS accent and one specifying an ‘Iberian’ accent. The speaker also identified individual sounds/word stress and sentence-level difficulties with one rater commenting that ‘mammal sounds like memo’. The speaker is actually describing the north of Kruger Park in South Africa at this point and says: ‘there are less mammals, more birds’ so this would appear to be a problem at the phoneme level with /æ/ in mammal pronounced as /ɛ/, a failure to use schwa /ə/ for the unstressed vowel and the non-standard elision of /l/ since ‘mammals’ /ˈmæməlz/ is pronounced [ˈmɛməʊz]. The raters also identified sentence-level and fluency difficulties, giving a median score of four for both, but the speaker only noted problems with the former. On the whole, however, there was speaker/rater agreement for this speaker.

Table 13. Speaker 11 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 11	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NNS accent	5/9
Influence on Accent	Portuguese	8 raters = NNS accent 1 rater = British
Individual sounds/ word stress difficulties	Yes	4/9
Sentence-level difficulties	No	5/9
Fluency difficulties	No	5/9

Speaker 11 described their accent as a ‘Portuguese accent in English’. Raters agreed that this was an NNS accent, although none specified a Portuguese accent. The speaker and rater also agreed that there were difficulties at the level of individual sounds/word stress, but raters also saw difficulties at the sentence-level and with fluency which were not noted by the speaker. There was, therefore, partial agreement between speaker and raters.

Table 14. Speaker 12 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 12	Speaker Self-Rating	Rater median rating on 9-point Likert scale
------------	---------------------	---

		(1=not NS/9=NS)
Nativeness	NNS accent	5/9
Influence on Accent	Portuguese	6 raters = NNS accent 1 rater = South American 1 rater = American 1 rater = I don't know
Individual sounds/word stress difficulties	Yes	6/9 1 rater = 'ad (had)
Sentence-level difficulties	Yes	5/9 1 rater = 'nice rhythm'
Fluency difficulties	No	6/9

Speaker 12 described their accent as a 'Portuguese accent in English'. The raters agreed that this was an NNS accent giving it a median score of five for nativeness and seven of the raters describing it as an NNS accent, with one specifying a 'South American' accent. The speaker identified individual sounds/word stress and sentence-level difficulties with one pointing out the omission of /h/ in the word 'had' /hæd/. While /h/ can be elided on weak forms of function words in standard British English, the elision occurred in the expression: 'we've had the chance', where 'had' is the main verb so elision would not occur. While the raters identified sentence-level difficulties with a median Likert score of five, one rater commented: 'nice rhythm' suggesting that they thought sentence stress was native-like. The raters scored moderately highly for fluency, with a median Likert score of six and this was not judged a problem area by the speaker. Overall, therefore, there is speaker/rater agreement for this speaker.

Table 15. Speaker 13 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 13	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NS accent	4/9
Influence on Accent	American English	4 raters = American 2 raters = NNS accent 2 raters = I don't know 1 rater = Iberian
Individual sounds/word stress difficulties	No	5/9 1 rater = 'pronunciation of d'
Sentence-level difficulties	No	6/9
Fluency difficulties	Yes	7/9 1 rater = 'some hesitancy'

Speaker 13 described their accent as an ‘American English’ accent. Raters did not agree giving the speaker a relatively low median mark of four for nativeness. However, four raters identified an American influence on the accent. Of the remaining raters, three identified an NNS accent with one specifying an ‘Iberian’ accent and the remaining two raters stating: ‘I don’t know’. Of the three measures, the speaker identified only fluency difficulties, but the raters gave a relatively high median score of seven, although one rater specified ‘some hesitancy’. Raters gave a lower median mark for individual sounds/word stress with a median score of five and one rater specifying ‘pronunciation of d’. The rater does not specify further and, despite listening to the token several times, it is unclear exactly what the rater meant by this. Raters also identified difficulties at the sentence-level not mentioned by the speaker. There is, therefore, only partial agreement between speaker and rater.

Table 16. Speaker 14 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 14	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NS accent	5/9
Influence on Accent	American English	5 raters = American 3 raters = NNS accent 1 rater = Iberian
Individual sounds/word stress difficulties	Yes	6/9 1 rater = ‘BERlin instead of BerLIN’
Sentence-level difficulties	No	5/9
Fluency difficulties	No	7/9 1 rater = ‘confident’

Speaker 14 described their accent as an ‘American English accent’. Raters did not agree, giving this speaker a median score of 5 for nativeness, although five raters detected an American influence. The other four raters described this as an NNS accent, with one rater specifying ‘Iberian’. Despite describing their accent as an NS accent, they identified individual sounds/word stress difficulties and this was in agreement with raters who gave a median score of 6 for this measure with one rater specifying a word-stress difficulty with ‘BERlin instead of BerLIN’. Raters gave this speaker a relatively high median score of 7 for fluency, with one rater describing them as ‘confident’, which was in broad agreement with the speaker who did not identify fluency as a difficulty. However, the speaker did not identify sentence-level features as a difficulty, and yet, this received a moderate median score of five. On the whole, there was partial agreement between speaker and raters for this speaker.

Table 17. Speaker 15 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 15	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	I don't know what I sound like	7/9
Influence on Accent	I don't know what I sound like	4 South African 4 British English 1 rater = NNS accent
Individual sounds/word stress difficulties	No	6/9 3 raters = distinctive oi in wife, while and I 1 rater = 'some odd vowels'
Sentence-level difficulties	No	7/9
Fluency difficulties	No	7/9

Speaker 15 stated: 'I don't know what I sound like in English' and in response to Question 8 asking which of the measures spoken English caused difficulty, the speaker ticked 'none of the above'. The raters gave a relatively high median score of seven for nativeness with four hearing a South African influence, four a British English influence and one stating this was an NNS accent. The speaker also received relatively high median score of seven for both sentence-level features and fluency, but a lower score of six for individual sounds/word stress difficulties. In particular, three of the raters noted the non-standard pronunciation of /aɪ/ in words such as 'wife' /waɪf/, 'while' /waɪl/ and 'I' /aɪ/, which one rater described as being pronounced as 'oi' (/ɔɪ/). Another rater noted: 'some odd vowels' but did not specify further. The raters gave this speaker a relatively high score for nativeness, but three raters highlighted the non-standard pronunciation of one phoneme in particular. However, the speaker stated that they did not know what they sounded like.

Table 18. Speaker 16 - Pre-Intervention comparison between speaker and rater accent judgements

Speaker 16	Speaker Self-Rating	Rater median rating on 9-point Likert scale (1=not NS/9=NS)
Nativeness	NS accent	6/9
Influence on Accent	American English	4 raters = British English 3 = I don't know 2 = NNS accent
Individual sounds/word stress difficulties	Yes	6/9 1 rater = 'odd pronunciation of r in mother'; 'de instead of the'

Sentence-level difficulties	No	6/9
Fluency difficulties	No	7/9

This speaker described their accent as an ‘American English accent’ with only word stress highlighted as a difficulty. Raters did not agree, giving it a moderate median score of 6 for nativeness, with four citing a British English influence, two classing this an NNS accent and three unable to classify the accent. Raters agreed with the speaker that there were difficulties at the level of individual sounds/word stress, with one rater specifying: ‘odd pronunciation of r in mother; de instead of the’. The former appears to be a reference to the speaker’s tendency to pronounce postvocalic /r/, which is more typical of standard American English (mother /'mʌðər/), as in standard British English this would be omitted (mother /'mʌðə/). The latter appears in the phrase ‘I visited places like the London Eye’, in which the /ðə/ is clearly pronounced [də]. The raters gave the speaker a relatively high median score of seven for fluency, but a lower score of six for sentence-level difficulties. Neither of these features were identified as problematic by the speaker. Overall, therefore, there is little agreement between speaker and raters for this accent.

Finally, it is important to mention that comments which were unclear were not included in the table. For example, Speaker 16 received the comment: “I don't know where she visited before big ben [sic]. Couldn't understand it.” It is unclear from the comment exactly where the communication breakdown occurred.

5.1.2 Summary of Findings

In view of the description of the individual speakers reported above, they seem to be broadly divided into 3 groups:

- 1) firstly, there are 5 speakers (3, 6, 7, 9 and 15) who stated that they did not know what they sounded like in English. This shows a certain level of honesty and self-reflection and belies the predictions of the Dunning-Kruger effect described in Section 2.4.1.4, which predicts discrepancy between how people judge their own abilities compared with how others judge them. In this case, the speakers were aware that they did not have an awareness of their own accent. It is to be noted, however, that the Dunning-Kruger effect posits that those with high ability levels are more circumspect about their achievements, but this would not appear to be the case here as only one speaker, Speaker 15, received a rating of 5 or above for nativeness among this group (Table 17).

- 2) there were seven speakers (1, 5, 8, 11, 13, 14 and 16) whose rating of their accent did not correspond to that of the native speaking raters in at least two measures. Four of these speakers (5, 13, 14 and 16) stated they had a NS American English accent. Table 19 gives a breakdown of individual rater judgements for these seven speakers:

Table 19. Pre-Intervention Mean Rater Score for Nativeness (1-9) for speakers rating their accent differently from NS raters (rounded to 1 decimal place)

Speaker	Mean rater Score for Nativeness	Standard Deviation	Range
Speaker 1	3.8	1.2	3-6
Speaker 5	4.9	1.4	4-8
Speaker 8	3.2	1.0	2-5
Speaker 11	3.2	1.0	2-5
Speaker 13	4.8	1.1	4-7
Speaker 14	5.3	1.7	3-7
Speaker 16	6.0	1.4	3-8

It is noticeable that there is a range of answers from the raters for each of these speakers and this appears to particularly so for speakers rated at above 5 for nativeness. Speakers 1, 8 and 11, who received relatively low mean nativeness rating of between 3.2 and 3.8 points, showed a rater standard deviation of between 1 and 1.2 points. However, speakers 5 and 13 who received moderate mean nativeness rating of 4.9 and 4.8 respectively, showed a rater standard deviation of 1.1 to 1.4, while speakers 14 and 16, who received the relatively high mean nativeness rating of 5.3 and 6 respectively, showed a rater standard deviation of 1.4 to 1.7 points. This suggests that there was greater rater agreement with speakers receiving mean rater judgements at the lower end of nativeness, while less rater agreement with speakers receiving mean rater judgements at the higher end of the nativeness scale. Speaker 16, for example, who received a mean nativeness rating of 6 points, received judgements of between 3 and 8 on the rating scale, meaning that while some raters evaluated them as having a strong NNS accent, others were judging this accent at close to nativeness. This finding suggests that different raters are evaluating these accents in quite different ways. It is also worth noting, however, that none of the raters gave a score of 9 to any of the speakers, but did give 9 virtually across the board to the NS tokens (speakers 25-27), which supports the findings of Chapter 1 suggesting that NSs are sensitive to the NNS-NS distinction (Section 1.2). The rater results for NS tokens will be discussed in Section 5.2.2 below.

- 3) the remaining 4 speakers (2, 4, 10, 12) showed an analysis of their pronunciation which broadly agreed with that of the NS raters. Speaker 2 was in complete agreement apart from the failure of the raters to specifically identify a Portuguese accent. Speaker 4 was in complete agreement apart from identifying fluency difficulties, which were rated at 7 out of 9 by the raters meaning the raters judged fluency to be approaching NS-level. Speakers 10 and 12 agreed with raters in all measures with the exception of fluency and NS raters did not specifically identify a Portuguese accent.

There were only 2 instances of speakers overestimating their difficulties: Speaker 4 and Speaker 13 judged themselves to have fluency difficulties, but both received a mean rating of 7 from the NS judges, meaning that they were judged to be approaching NS level.

5.1.3 Discussion – Question 1

It is important to state that the speakers and the raters were asked to analyse the accents in different ways. The speakers were asked to identify their accent and give a yes/no answer as to whether they had problems in specific areas. The raters were also asked to identify the accent, but were then asked to judge each measure, including ‘nativeness’ on a scale of 1-9. It is difficult to claim that a description of accent and a judgement of nativeness are comparable. In order for the two ratings - that of the speaker and of the raters - to be comparable, they should have been judged using the same method. At the time these ratings were carried out, it was felt that asking participants to complete the Likert scale ratings of their own accent would be unproductive since, asking them to listen to their speech token and rate it would already be raising their awareness to their own accent. The intention was to get an idea of how they perceived their L2 accent prior to any intervention. A possible solution would have been to ask the participants to complete the rater questionnaire regarding their own accent, but without asking them to listen to their speech token. This was a limitation of this part of the research, but nevertheless it was possible to extrapolate a certain amount of information from the data obtained.

It was noticeable that, in many cases, there was a marked difference of opinion between the raters. For example, Speaker 15 was rated as having a South African influenced accent by four raters, but as having a British influenced English accent by four other raters. As explained in Section 5.1.1, the raters were deliberately chosen to provide variety: American/British; familiarity with Portuguese accent/no familiarity; experience in language teaching or linguistics/no experience and ages ranging from 25 to 71 at the start of the study. This means that the raters were not a homogeneous group and would appear

to be responding differently to the different speakers. More needs to be done to research the influence of NS rater background on rating scores, but that is beyond the scope of this research.

Six speakers stated they had a Portuguese accent in English. Raters agreed that these speakers were NNS giving them a mean score of between 2/9 and 5/9 for nativeness, but no raters specifically mentioned a Portuguese accent, despite the fact that five of the nine were familiar with the accent. The raters were given no information regarding the speakers at any point. It is possible that these raters were so familiar with this accent that they did not think to include this information. 'Portuguese accent in English' was not given as one of the options as it was felt this would 'lead' the raters, but this was possibly a mistake. Perhaps a more general question such as, 'how would you describe the accent?' would have elicited more information in this area.

Furthermore, it could also be argued that the participants were not representative of a typical language class if, indeed, there is such a thing. The fact that they chose to take part in a pronunciation course suggests that they were, for whatever reason, unhappy with their pronunciation at the start of the course and implies a certain level of self-awareness of their own accent. Also, while information about their employment was not included in the Learner Consent Form and Pronunciation Questionnaire (Appendix D), it became apparent that several of the speakers were either teachers or worked in the area of second language acquisition. It might be expected, therefore, that these participants would be more informed about pronunciation than those with no experience in the area.

Overall, these findings suggest that that phonological self-awareness, as proposed in Section 2.4.1.4, is an ID variable - that some people are self-aware to a greater extent and others to a lesser extent. This provides support for the idea that it can be considered an element of Language Aptitude.

5.1.4 Conclusion – Question 1

Question 1 asked: *To what extent do learners' pre-intervention assessments of their own spontaneous speech production correspond to those of NS raters?* and it was predicted that, in line with the predictions of the Dunning-Kruger effect (Section 2.4.1.4), many learners would judge their accents more favourably than would NS raters. This was borne out by the results, although there was a certain amount of individual variation in learner phonological self-awareness. In this cohort, 25% of the speakers produced judgments of their accent which were close to that of the NS raters, but the remaining 75% were either unable to describe their accent or described it differently from the NS raters. As previously mentioned, the fact that the speakers chose to take part in a pronunciation course and that several of them worked in the area of language acquisition suggests that this cohort would, in fact, be *more* self-aware than the average learner, which would imply that the actual percentage of

learners who are able to match NS judgements of their accent may be even smaller than the 25% identified here.

Clearly a much larger cohort would be needed in order to judge whether this figure is in any way representative of the population as a whole and results may have been influenced by methodological problems and rater variation as mentioned above. However, two broad conclusions can be drawn. Firstly, it provides some support for the notion that many students are unaware of how others judge their L2 accent, but more research is needed in order to delve more deeply into this finding. Secondly, if it is the case that many learners lack awareness of their own accent, raising awareness surely has to be the starting point of any proposed intervention as suggested in Section 3.2.2 (Couper, 2015; McGregor & Reed, 2018). It is difficult to imagine how students can make any coherent decisions regarding their pronunciation goals, if they are unable to accurately evaluate their own output and are unaware of how they are being perceived by listeners.

5.2 QUESTION 2: WILL EITHER OR BOTH OF THE INTERVENTIONS TO INFLUENCE PRONUNCIATION PRODUCE SUPERIOR RESULTS TO MERE EXPOSURE TO A NATIVE SPEAKER ACCENT IN LATE LEARNERS?

To answer this question, the rater judgements of the two intervention groups, Group 1, who received bottom-up instruction and Group 2, who received top-down instruction, were compared with rater judgements of the Control, who attended ESL classes delivered by an NS of British English without PI, at the three time points: Pre-, Post- and Delayed Post-Intervention. It was hoped to ascertain whether any of the three conditions produced changes in pronunciation in the direction of nativeness according to the rater judgements. The data was analysed using descriptive statistics to identify any obvious trends and further analysed using inferential statistics in the form of a mixed ANOVA (Analysis of Variance) to identify whether any broader generalisations could be made about the data. It was predicted that, in agreement with much previous PI research as described in Sections 3.3 and 3.4, explicit intervention (Groups 1 and 2) would produce positive outcomes, while mere exposure (Control) would have no effect on pronunciation (Lee et al., 2015; Saito, 2012; Thomson & Derwing, 2015a).

5.2.1 Inter-rater Reliability

In order to ascertain the degree of homogeneity in the rater scores, inter-rater reliability scores (Cronbach's alpha) were calculated using SPSS for each of the six measures individually (nativeness,

sounds, sentence-level features, fluency, vocabulary and accuracy). The results for NSs were omitted from the calculation as there was minimal variation in the rater score (see Section 5.2.2 below). The results were as follows: nativeness ($\alpha = .868$), segmental/word-level features ($\alpha = .826$), sentence-level ($\alpha = .803$), fluency ($\alpha = .868$), vocabulary ($\alpha = .755$) and accuracy ($\alpha = .733$). This means that reliability for nativeness, segmental/word-level features, sentence-level features and fluency is very good (0.81-0.90), for vocabulary and accuracy is good (0.71-0.80) (Konting et al., 2009, George & Mallery, 2003). It can be concluded that there was broad agreement between the raters for every measure.

5.2.2 Native Speaker Scores

Three British native speakers of English (Speakers 25-27) were included randomly among the NNS tokens in order to determine whether the raters were able to assess them as such. Figure 1 shows that the native speakers were given high ratings at each of the three time points: Pre-Intervention (Pre), Post-Intervention (Post) and Delayed Post-Intervention (Delayed) for an overall mean score of the six measures, in other words, nativeness, sounds, sentence-level features, fluency, vocabulary and accuracy added together.

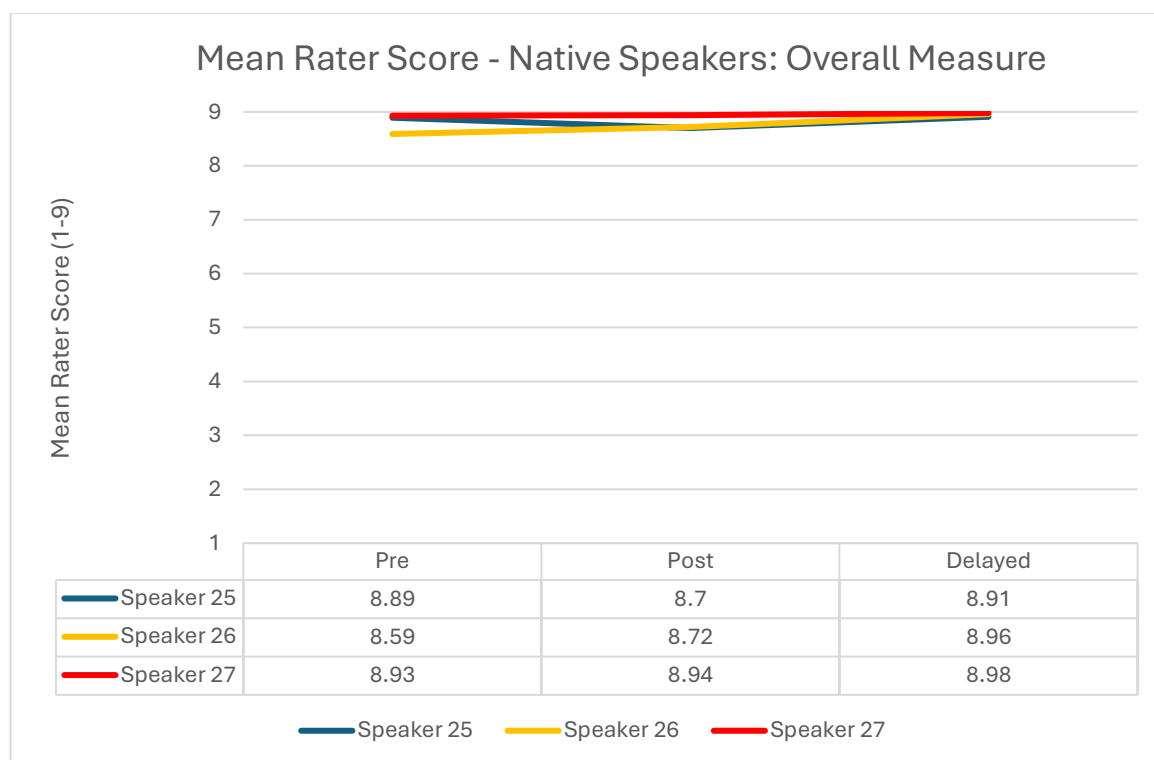


Figure 1. Overall Mean Rater Score: Native Speakers (rounded to 2 decimal places)

When these figures are broken down into individual measures (Figure 2), it can be seen that the overall picture remains the same with a slight dip in attainment for fluency. This is to be expected given that the speakers were asked to give spontaneous speech samples, which had not been prepared.

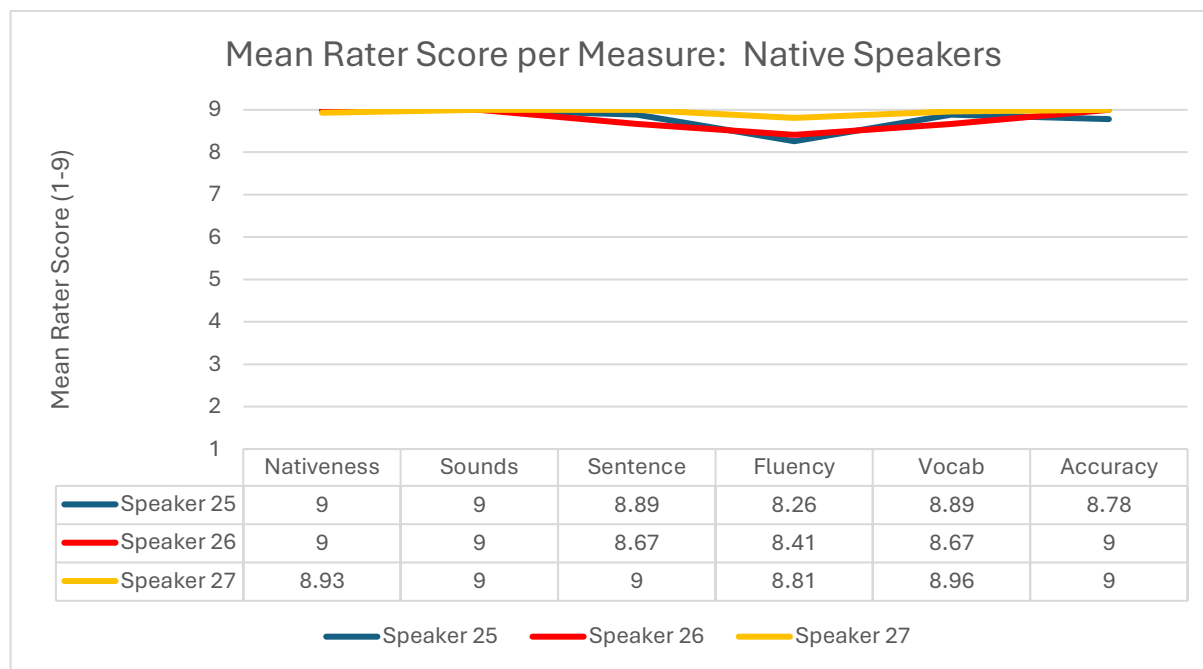


Figure 2. Mean Rater Score per Measure: Native Speakers (rounded to 2 decimal places)

The native speakers were always given 9 for nativeness and sound-level features, with the exception of Speaker 27, who received two 8s for nativeness in the Delayed condition. Speaker 27 spoke SBE, but did have some features of a regional British accent, so it is possible that the American raters did not recognise this accent and judged it accordingly. However, as the ratings were anonymous, this is speculation. The overall picture shows that the raters gave very high scores to the native speakers for all measures and were judged as native speakers or near-native speakers at all three time points. Given that purpose of including the NS tokens was to ensure they were judged as such, these results prove this to be the case and their results were not included in the analysis beyond this point.

5.2.3 Analysis of Overall Measure (Six Measures Combined) Group 1, 2 and Control

The Overall Measure gave the scores for nativeness, sounds, sentence-level features, fluency, vocabulary and accuracy added together. A comparison of mean scores for each group shows very little difference between the groups at each time point (Figure 3). There was a general tendency towards an increase in rater judgement scores towards nativeness, with Group 1 showing a 1.1-point increase from Pre- to Delayed, Group 2 showing a 0.4 point increase and the Control increasing by 0.2 of a point.

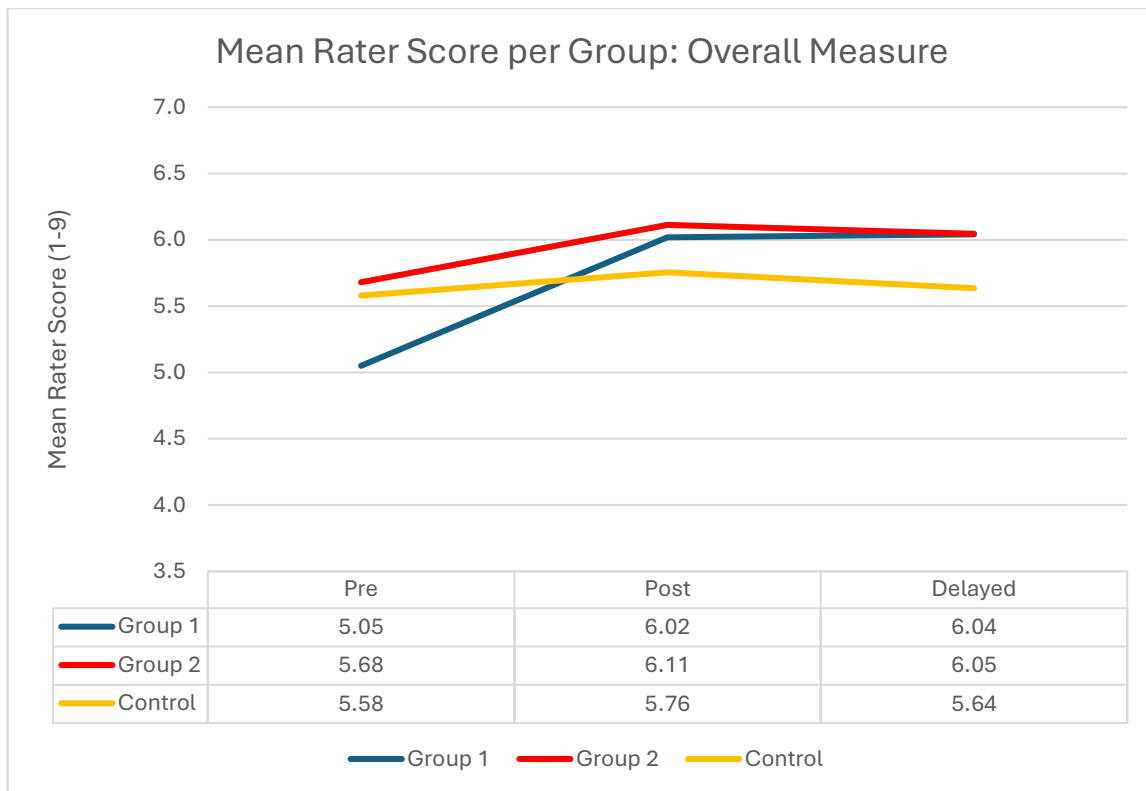


Figure 3. Mean Rater Score per Group: Overall Measure (rounded to 2 decimal places) – Group 1, Group 2 and Control

When the figures are broken down into individual speakers, a more nuanced picture can be seen (Table 20). Group 1 saw an increase for every speaker except Speaker 4 (-0.34) from Pre- to Delayed, while in Group 2 only Speaker 15 (-1.48) did not increase. In the Control group, 5 speakers saw a drop in rating from Pre- to Delayed (Speakers 17, 18, 22, 23 and 24) with only 3 speakers showing an increase (Speakers 19, 20 and 21).

Table 20. Mean Rater Score per Speaker: Overall Measure - Group 1, Group 2 and Control (rounded to 2 decimal places)

Group	Speaker	Mean Rater Score: Pre	Mean Rater Score: Post	Mean Rater Score: Delayed	Mean change in Rater Score Pre-Delayed
1	1	5.13	5.81	5.83	0.70
	2	3.89	5.15	5.26	1.37
	3	3.85	6.24	5.85	2.00
	4	6.65	7.07	6.31	-0.34
	5	5.48	6.39	5.87	0.39
	6	5.55	5.96	6.80	1.25
	7	5.67	6.20	6.22	0.55
	8	4.20	5.37	6.19	1.99
	MEAN	5.05	6.02	6.04	0.99
2	9	4.65	5.61	6.20	1.56
	10	4.02	6.11	4.85	0.83
	11	5.19	5.28	5.59	0.41
	12	5.85	6.43	6.30	0.44
	13	6.04	6.69	6.44	0.41
	14	6.22	5.94	6.43	0.20
	15	7.00	6.04	5.52	-1.48
	16	6.44	6.80	7.04	0.59
	MEAN	5.68	6.11	6.05	0.37
Control	17	5.19	4.19	4.57	-0.62
	18	6.26	6.24	6.07	-0.19
	19	4.94	7.28	5.96	1.02
	20	6.15	6.76	6.83	0.68
	21	5.96	6.26	6.67	0.74
	22	4.52	4.22	4.43	-0.09
	23	6.85	6.85	6.09	-0.76
	24	4.76	4.24	4.46	-0.30
	MEAN	5.58	5.76	5.64	0.06

The mean rater scores were then analysed for inferential findings. An ANOVA test was chosen as this test has been used by other researchers to evaluate Likert-scale data (Saito et al., 2016; Galante & Thomson, 2017; Foote & McDonough, 2017). A 3x(3) mixed ANOVA was chosen as the data included both a between-groups variable (Group 1, Group 2 and Control) and a within-groups variable (Pre-, Post- and Delayed Post-Test). All ANOVA scores were calculated using SPSS. There were statistically significant effects of Time $F(2,42) = 7.01$, $p = <0.002$, but not for Group $F(2,21) = 0.71$, $p = 0.32$ and no significant interaction between Time and Group $F(4,42) = 1.83$, $p = 0.14$ (Table 21).

Table 21. Mixed ANOVA: Overall Rater Score - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	4.04	2	2.08	7.01	<0.002
Time x Group	2.10	4	0.53	1.83	0.14
Error (Time)	12.07	42	0.29		
Group	1.14	2	0.57	0.71	0.32
Error (Group)	34.78	21	1.66		

These inferential results back up the descriptive findings that time was a significant factor, but group was not significant and there was no significant interaction between time and group; in other words, the overall trend was for learners to achieve higher mean rater scores over time regardless of the group they were in.

That there was no difference between the Control and the intervention groups is not as surprising as it might seem given that the overall scores included measures which were not explicitly addressed in the intervention, namely fluency, vocabulary and accuracy. In order to shed more light on this finding, each measure (nativeness, sounds, sentence-level features, fluency, vocabulary, accuracy) was analysed separately.

5.2.4 Analysis of Measures Individually (Group 1, 2 and Control)

5.2.4.1 Nativeness

A comparison of mean rater scores for nativeness at the three time points (Figure 4) shows that Group 1 was initially judged to be the least ‘native’ group and the Control was marginally the stronger group. However, Post-Intervention, the two intervention groups overtook the Control and maintained this advantage in the Delayed Post-Intervention. Group 1 showed a mean change of 1.52 points from Pre- to Delayed and Group 2 showed an increase of 0.86 of a point. The Control group showed almost no change between Pre and Delayed with only a 0.05 of a point increase.

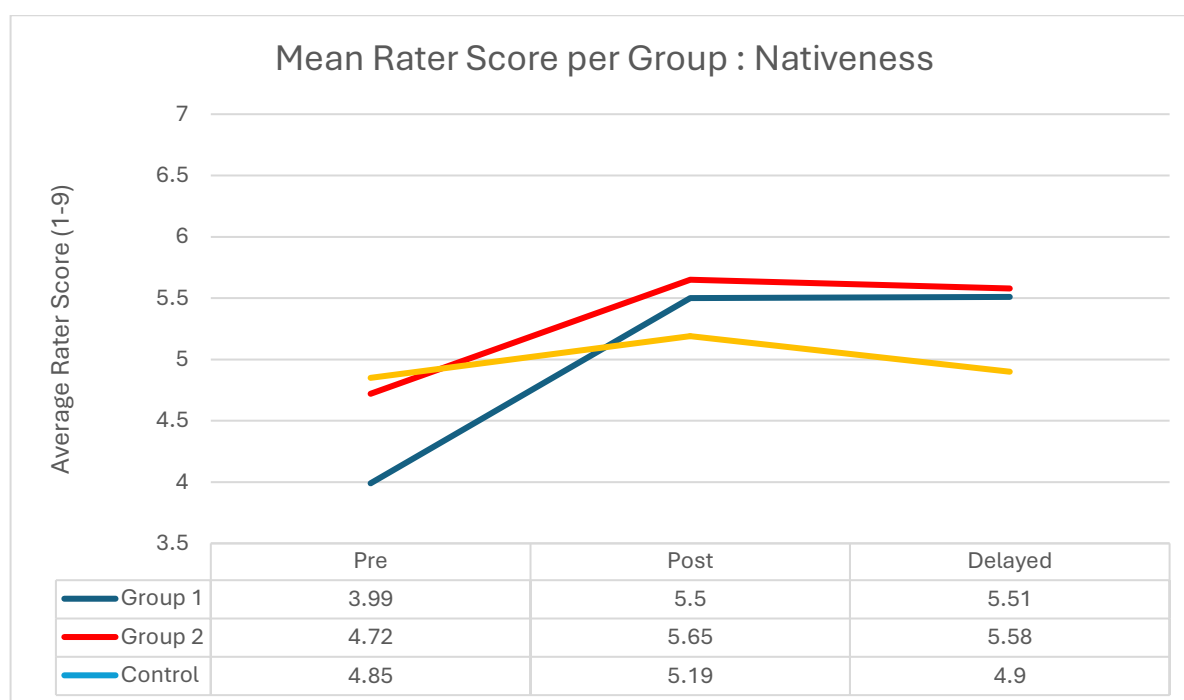


Figure 4. Mean Rating Score per Group: Nativeness (rounded to 2 decimal places) – Group 1, Group 2 and Control

Looking at mean rater scores for individual speakers (Table 22) shows that all the speakers in Group 1 and seven of the eight in Group 2 achieved change in the direction of nativeness according to the raters. However, in the Control Group three speakers (22, 23 and 24) saw a decrease in their mean rater score for nativeness. The exception in Group 2 was Speaker 15 who saw a drop of 1.34 points from Pre- to Delayed Post-Intervention. By way of an explanation for the atypical results obtained for Speaker 15, it should be noted that they had lived in SE England for some time and their English contained features of a regional British accent, most noticeable of which was the pronunciation of the standard /aɪ/ as [ɔɪ] for example: *my wife* /maɪ waɪf/ was pronounced [mɔɪ wɔɪf]. However, they also showed non-standard pronunciation of features *not* present in this regional accent such as /ð/ pronounced [d], for example *they* /ðeɪ/ was pronounced [deɪ]. This apparent contradiction seems to have confused raters. The specific sounds described here were mentioned by several of the raters with the speech samples produced by this speaker being described as ‘odd’ on three separate occasions by raters and his accent being variously judged as influenced by Australian, British, South African and New Zealand English. This apparent confusion on the part of the raters may have led to the atypical results received for this speaker. It is also worth noting that this speaker stated in the End of Course Questionnaire that they were happy with their current accent and so was, presumably, not motivated to modify it in any way.

Table 22. Mean Rater Score per Speaker: Nativeness - Group 1, Group 2 and Control (rounded to 2 decimal places)

Group	Speaker	Mean Rater Score: Pre	Mean Rater Score: Post	Mean Rater Score: Delayed	Mean change in Rater Score Pre-Delayed
1	1	3.78	4.67	5.11	1.33
	2	2.44	4.78	4.89	2.45
	3	2.89	5.67	5.22	2.33
	4	5.56	6.56	5.67	0.11
	5	4.89	6.00	6.33	1.44
	6	4.89	5.56	6.44	1.55
	7	4.22	6.11	5.22	1.00
	8	3.22	4.67	5.22	2.00
	MEAN	3.99	5.50	5.51	1.53
2	9	4.22	5.44	5.78	1.56
	10	2.56	5.33	4.11	1.55
	11	3.22	4.33	4.78	1.56
	12	5.11	5.56	5.67	0.56
	13	4.78	5.78	5.56	0.78
	14	5.33	5.78	6.33	1.00
	15	6.56	6.44	5.22	-1.34
	16	6.00	6.56	7.22	1.22
	MEAN	4.72	5.65	5.58	0.86
Control	17	4.00	3.33	4.11	0.11
	18	5.00	5.33	5.11	0.11
	19	4.00	6.89	5.56	1.56
	20	5.67	6.67	6.56	0.89
	21	4.67	5.44	6.00	1.33
	22	4.44	3.78	3.33	-1.11
	23	6.78	6.67	5.33	-1.45
	24	4.22	3.44	3.22	-1.00
	MEAN	4.85	5.19	4.90	0.06

The nativeness scores were further analysed using a 3x(3) mixed ANOVA (Table 23) with the between-subjects factor of Group (Group 1, Group 2, Control) and the within-subjects factor of Time (Pre-, Post-, Delayed-Intervention). There were statistically significant effects for Time $F(2,42) = 15.22$, $p = <0.001$, but not for Group $F(2,21) = 0.35$, $p = 0.711$. However, there was significant interaction between Time and Group $F(4,42) = 2.99$, $p = 0.029$.

Table 23. Mixed ANOVA: Nativeness - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares	Degrees of Freedom	Mean Square	F ratio	p-value
Time	12.34	2	6.17	15.22	<0.001
Time x Group	4.85	4	1.21	2.99	0.029
Error (Time)	17.02	42	0.41		
Group	1.73	2	0.87	0.35	0.71
Error (Group)	52.58	21	2.50		

This means that the change over time was significant, but also which group the participants were in was affecting the results at the three time points. While the ANOVA does not indicate the direction of the influence, it can be seen from Figure 4 above that the change was positive for Groups 1 and 2, but neutral for the Control. This means that, in the case of nativeness, both interventions were more effective than mere exposure in provoking an increase in rater scores.

5.2.4.2 Segmental/Word-Level Features

A comparison of mean rater scores for segmental/word-level features at the three time points (Figure 5) shows that Group 1 was initially judged to be the least native in terms of segmental/word-level features. Group 2 and the Control were judged to be at a similar level. However, Post-Intervention, the two intervention groups overtook the Control and maintained this advantage in the Delayed Post-Intervention. Group 1 increased by 1.14 point from Pre- to Delayed, while Group 2 increased by 0.83 of a point. The Control showed almost no change (0.04) between Pre- and Delayed Post-Test.

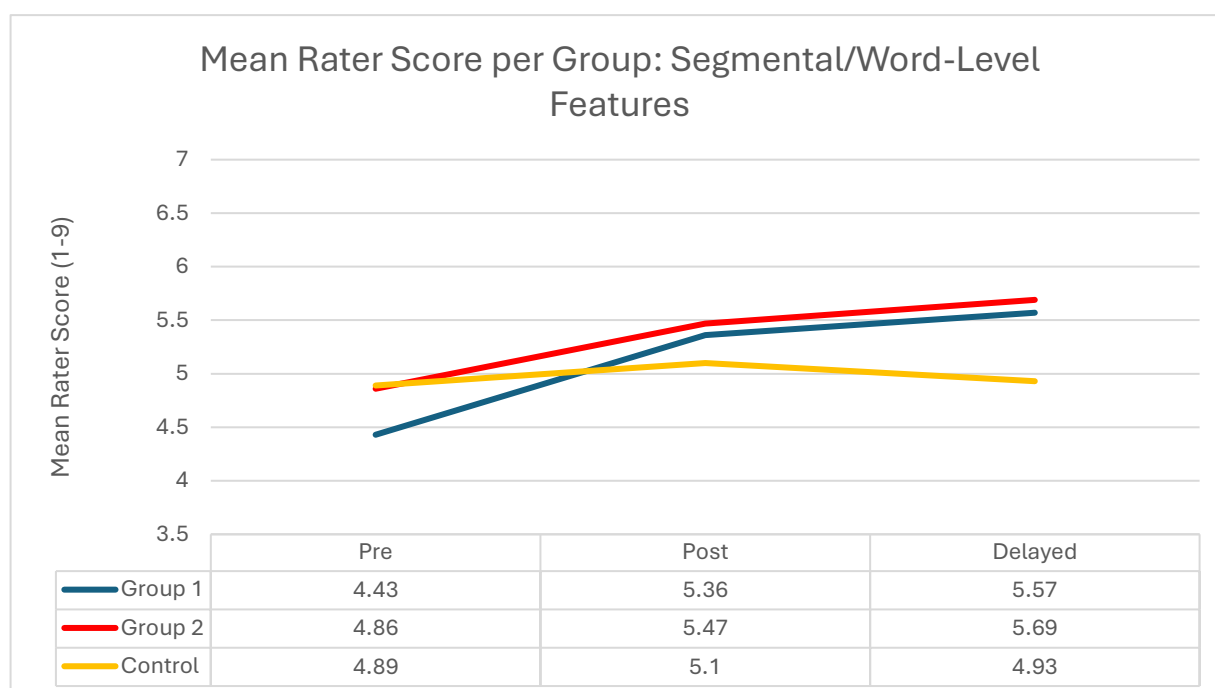


Figure 5. Mean Rater Score per Group: Segmental/Word-Level Features (rounded to 2 decimal places) – Group 1, Group 2 and Control

The segmental/word-level scores were further analysed using a 3x(3) mixed ANOVA. There were statistically significant effects for Time $F(2,42) = 9.47$, $p = <0.001$, but not for Group $F(2,21) = 0.36$, $p = 0.704$ and no significant interaction between Time and Group $F(4,42) = 1.98$, $p = 0.115$ (Table 24).

Table 24. Mixed ANOVA: Segmental/Word-level Features - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	6.39	2	3.20	9.47	<0.001
Time x Group	2.67	4	0.67	1.98	0.115
Error (Time)	14.16	42	0.34		
Group	1.67	2	0.83	0.36	0.704
Error (Group)	49.07	21	2.34		

This means that the change over time was significant but, contrary to the evidence of the descriptive statistics, which group a participant was in was not affecting the results at the three time points, so the intervention was not significantly more effective than mere exposure in achieving an increase in rater scores for segmental/word-level features.

5.2.4.3 Sentence-Level Features

An analysis of the mean rater scores at the three time points for sentence-level features shows that, similar to the results for segmental/word-level features, the two intervention groups showed an increase in mean rater score, while the Control did not (Figure 6). In fact, Pre- to Delayed, the Control showed a drop of 0.06 of a point, while Group 1 increased by 1.14 points and Group 2 by 0.42 of a point.

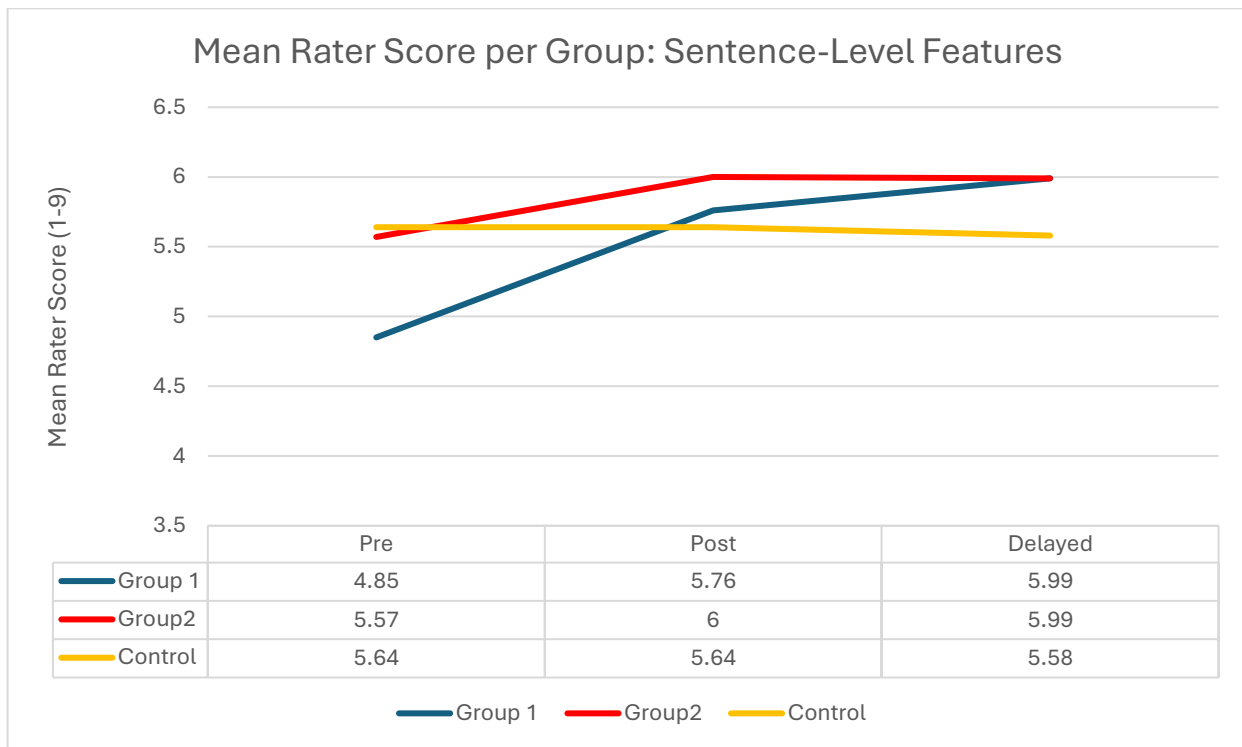


Figure 6. Mean Rater Score per Group: Sentence-Level Features (rounded to 2 decimal places) – Group 1, Group 2 and Control

The sentence-level scores were further analysed using a 3x(3) mixed ANOVA. There were statistically significant effects for Time $F(2,42) = 6.18$, $p = .004$, but not for Group $F(2,21) = 0.35$, $p = 0.71$. However, there was significant interaction between Time and Group $F(4,42) = 2.70$, $p = 0.043$ (Table 25).

Table 25. Mixed ANOVA: Sentence-level Features - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	3.64	2	1.82	6.18	.004
Time x Group	3.18	4	0.79	2.70	0.043
Error (Time)	12.36	42	0.29		
Group	1.31	2	0.65	0.35	0.71
Error (Group)	39.37	21	1.87		

The change over time was significant and there was a significant interaction between the group a participant was in and the change obtained. The descriptive statistics indicate that this was due to the Control Group remaining largely stable across the three time points, while Group 1 and Group 2 showed an increase. In short, both intervention groups showed an increase in mean rater score, while the Control Group did not and the difference in attainment was statistically significant.

5.2.4.4 Fluency

Analysed for fluency, the statistics show a decidedly different picture (Figure 7). Group 2 and the Controls behave in a similar way over the three time points, Group 2 showing a minimal decline of -0.08 of a point and the Control declining by -0.03 from Pre to Delayed. However, Group 1 shows a marked increase of 1.10 points from Pre to Delayed.

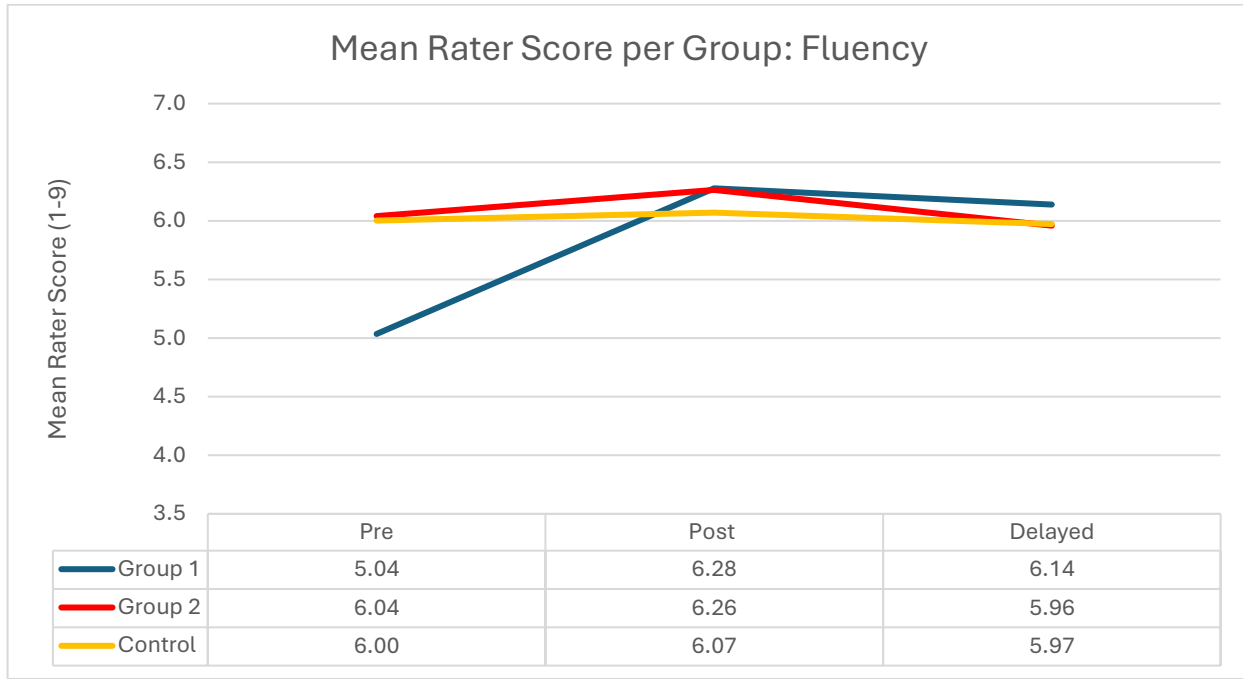


Figure 7. Mean Rater Score per Group: Fluency (rounded to 2 decimal places) – Group 1, Group 2 and Control

The fluency scores were further analysed using a 3x(3) mixed ANOVA. There were no statistically significant effects for Time $F(2,42) = 3.12$, $p = 0.054$, or for Group $F(2,21) = 0.22$, $p = 0.803$ and no significant interaction between Time and Group $F(4,42) = 2.24$, $p = 0.081$ (Table 26).

Table 26. Mixed ANOVA: Fluency - Group 1, Group 2 and Control - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	3.23	2	1.61	3.12	0.054
Time x Group	4.63	4	1.16	2.24	0.081
Error (Time)	21.69	42	0.52		
Group	0.94	2	0.47	0.22	0.803
Error (Group)	44.61	21	2.12		

Therefore, despite the apparent superiority of Group 1 in this measure, the findings were not statistically significant. There was no significant increase over time and no significant interaction between time and group. This finding is discussed further in the Section 5.2.5 below.

5.2.4.5 Vocabulary

The change from Pre- to Delayed is relatively limited for the vocabulary measure (Figure 8). Group 1 showed an increase of 0.7 of a point from Pre- to Delayed, Group 2 increased by 0.04 and the Control increased by 0.12.

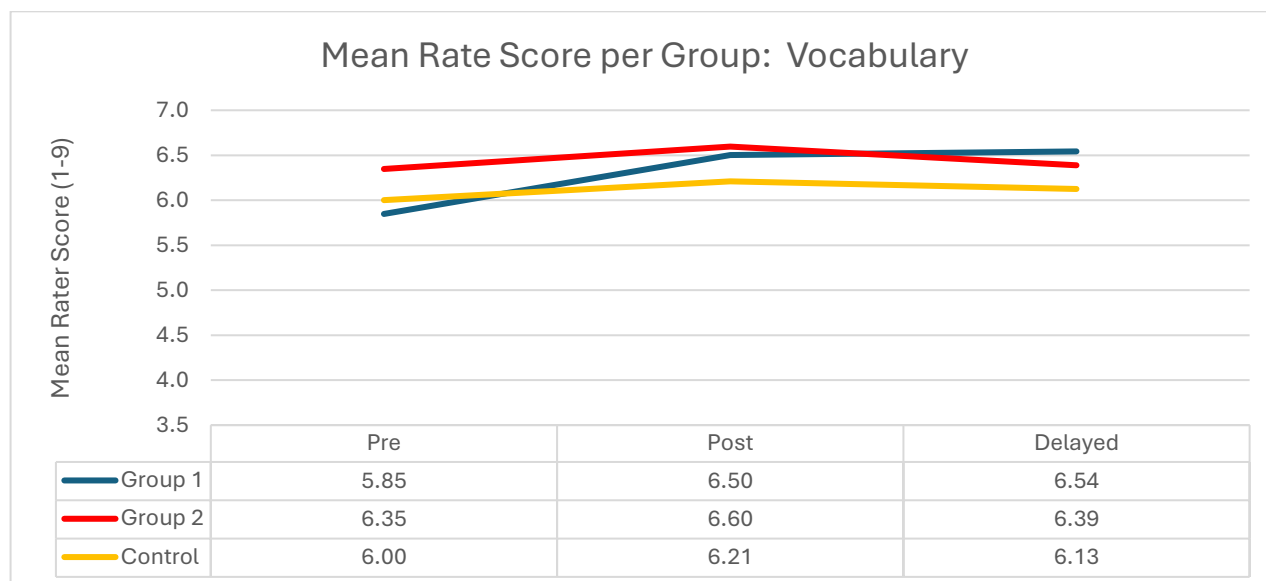


Figure 8. Mean Rater Score per Group: Vocabulary (rounded to 2 decimal places) – Group 1, Group 2 and Control

The vocabulary scores were further analysed using a 3x(3) mixed ANOVA. There were no statistically significant effects for Time $F(2,42) = 2.56$, $p = 0.09$, or for Group $F(2,21) = 0.58$, $p = 0.569$ and no significant interaction between Time and Group $F(4,42) = 0.76$, $p = 0.555$ (Table 27).

Table 27. Mixed ANOVA: Mean Scores for Vocabulary - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	1.81	2	0.91	2.56	0.09
Time x Group	1.08	4	0.27	0.76	0.555
Error (Time)	14.88	42	0.35		
Group	1.33	2	0.66	0.58	0.569
Error (Group)	24.08	21	1.15		

For vocabulary, there were no significant results for time, group or the interaction between time and group meaning that vocabulary was not influenced by the intervention. This is to be expected since vocabulary expansion was not addressed in the intervention.

5.2.4.6 Accuracy

Accuracy was the measure which showed the least change over the three time points. From Pre- to Delayed, Group 1 increased by 0.31 of a point, Group 2 by 0.15 and the Control by 0.22 (Figure 9).

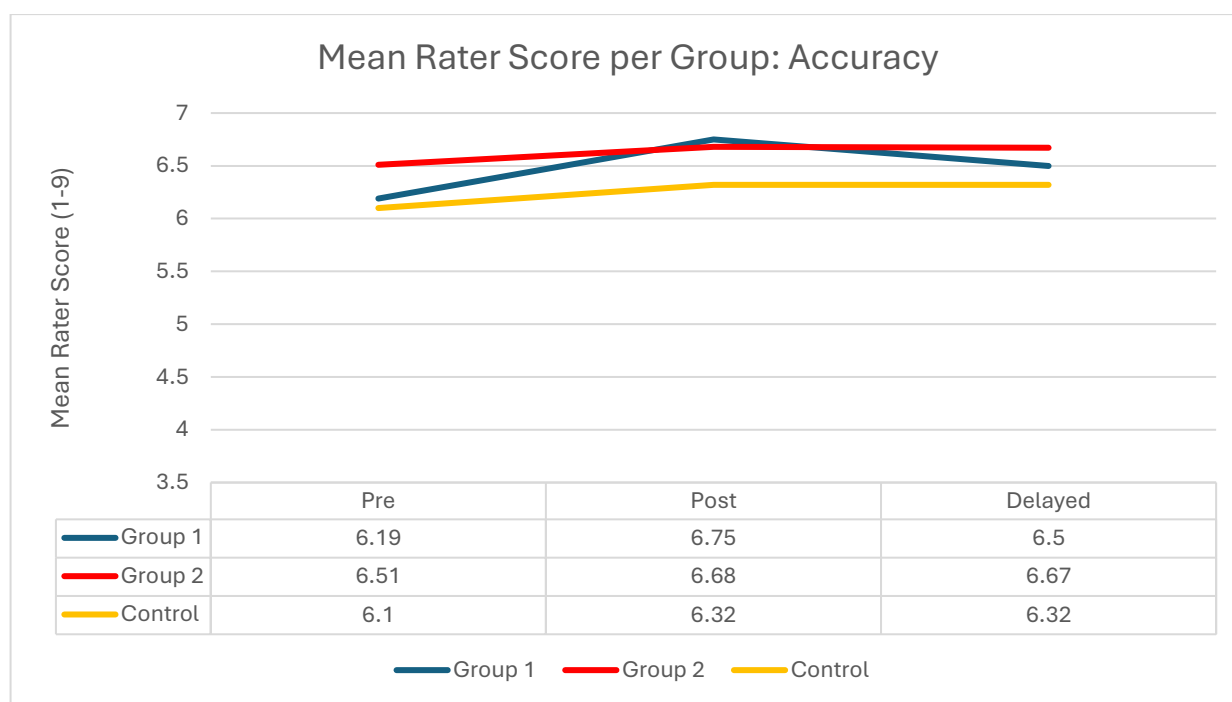


Figure 9. Mean Rater Score per Group: Accuracy (rounded to 2 decimal places) – Group 1, Group 2 and Control

The accuracy scores were further analysed using a 3x(3) mixed ANOVA. There were no statistically significant effects for Time $F(2,42) = 2.56$, $p = 0.09$, or for group $F(2,21) = 0.58$, $p = 0.569$ and no significant interaction between Time and Group $F(4,42) = 0.76$, $p = 0.555$ (Table 28).

Table 28. Mixed ANOVA: Accuracy - Group 1, Group 2 and Control (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	1.26	2	0.63	2.16	0.127
Time x Group	0.37	4	0.09	0.32	0.862
Error (Time)	12.26	42	0.29		
Group	1.72	2	0.86	0.64	0.536
Error (Group)	28.36	21	1.35		

For accuracy, there were no significant results for time, group or the interaction between time and group meaning that accuracy was not influenced by the intervention. This is to be expected since language accuracy was not addressed by the intervention.

5.2.5 Discussion – Question 2

In the case of nativeness and sentence-level features there was a significant finding for time and an interaction effect meaning that the effect of time on rating score was influenced by which group the participant was in. In Figure 4 and Figure 6, it can be seen that while the Control remained relatively unchanged at the three time points, the two intervention groups showed an increase. This suggests that, while the intervention was helping Groups 1 and 2 to achieve higher nativeness ratings from the raters, this was not the case with the Control, whose production remained largely unchanged for these two measures. This supports previous findings suggesting that exposure alone may be insufficient to promote L2 pronunciation acquisition in late learners (Schmidt, 2012) and that PI is effective in achieving accent modification (Kennedy & Trofimovich, Lee et al., 2015; Saito, 2012; Thomson & Derwing, 2015a).

For Segmental/Word-level features, time is shown to be significant, but the interaction between time and group is not significant, meaning there was change over time, but that this was not influenced by which group the participant was in. This seems to contradict the descriptive statistics shown in Figure 5, where the intervention groups appear to show an increase, while the Control does not. In fact, the lack of a significant finding here is probably due to the fact that all three groups show an increase between Pre- and Delayed Post-Test. However, while the change in Group 1 is 1.14 and Group 2 is 0.83 (Figure 5), the change in the Control is only 0.04, which shows that the change in the intervention groups was much greater than the change in the Control. Possibly, a larger sample size would have achieved a significant result for this measure.

The most surprising finding is the Group 1 jump in Fluency from Pre- to Post-Test. Although this was found not to be statistically significant, it is apparent in the descriptive analysis. In Figure 3, it can be seen that Group 1 was initially judged by the raters to be the weaker group and this was also noticeable in the participants self-reported level of English (Section 5.3 above). In Group 2, two of the participants stated their level as C2 and one who did not know the level was an English teacher who had lived in the UK. None of the participants of Group 1 claimed C2 level. This could mean that regular English language interaction with NS English had a greater effect on Group 1 than on Group 2, several of whom already had a high level of English. However, this does not explain why there was no

change in fluency for the Control, who were all assessed as B2 level in English and who were involved in English language interactions with an English NS in the classroom.

This could be a feature of age differences between the groups. In Table 2 in Section 4.3 shows the age of participants at the start of the course and the mean age of each group can be seen below in Table 29.

Table 29. Age of Participants at Start of Study

Group	Age of Participants								Mean Age
Group 1	49	28	51	21	23	18	18	41	31.1
Group 2	18	28	28	27	18	18	41	20	24.8
Control	32	19	38	20	18	23	19	18	23.3

It shows that while Group 2 and the Control were similar in age (24.8 years and 23.3 years respectively), the mean age of Group 1 was somewhat higher at 31.1 years at the start of the course. Younger Portuguese learners of English are increasingly exposed to NS spoken English via streaming platforms such as Netflix and Disney+ and via voice-chat on online multiplayer computer games such as Call of Duty and Fortnite. Student level is usually assessed for productive (speaking and writing) and receptive skills (reading and listening). This is true of major examining boards such as IELTS and Cambridge, which many of the participants used as proof of their level. This is merely speculation, but it is possible that the Control Group, being younger, had a level of spoken English above B2, while their written English was at or below B2, giving them an overall rating of B2. The older group, Group 1, probably did not have this level of exposure to NS English. Therefore, while Group 2 and the Control already had a good level of fluency, the older group, Group 1, were less fluent and therefore possibly able to make greater gains from regular interactions in the classroom. Future research should address this issue by balancing the groups for age.

For the final two measures, vocabulary and accuracy, Figures 8 and 9 respectively show that there was little change in the groups over the three time periods and the inferential statistics (Tables 27 and 28) show that no significant results were obtained. As mentioned in the analysis, this result is to be expected since these measures were not targeted in the interventions.

This finding needs to take account of the fact that the sample sizes were relatively small with just eight participants in each group and that this may have limited the scope of the findings. As discussed in Section 5.3 above, several participants did not start the course and others were dropped from the study

as they did not complete a minimum of 80% of the course, which depleted the numbers involved. It is important, going forward, to ascertain whether these results would be obtained with a larger cohort of participants.

5.2.6 Conclusion – Question 2

In conclusion, for the second research question, “*Will either or both of the interventions to influence pronunciation produce superior results to mere exposure to a native speaker accent in late learners?*” the answer is, broadly speaking, ‘yes’. This finding supports the prediction that, in agreement with previous research, pronunciation intervention will prove more effective than mere exposure (Lee et al., 2015; Saito, 2012; Thomson & Derwing, 2015a). A significant interaction was found between time and group for nativeness and this appeared to be largely due to a change in sentence-level features, which also showed a statistically significant interaction between time and group. For segmental/word-level features, there was also greater increase in the intervention groups than in the Control, although this was not statistically significant. For fluency, Group 1 was weaker than the other two groups in the Pre-Intervention Condition, but caught up in the Post-Intervention and Delayed Post-Intervention, although the difference was not significant. The behaviour of Group 1 was possibly due to the fact that they, being initially the weaker group, had more room to improve as a result of the regular contact with English provided by the intervention. No statistically significant difference between the groups was found for vocabulary and accuracy, which is to be expected as these areas were not the focus of the intervention.

5. 3 QUESTION 3: WILL A TOP-DOWN STRATEGY-LED APPROACH BRING LATE LEARNERS CLOSER TO ACHIEVING A NATIVE-LIKE ACCENT THAN A BOTTOM-UP STRUCTURE-LED APPROACH?

In order to answer this question, the rater judgements of the two intervention groups, Group 1, who received bottom-up instruction and Group 2, who received top-down instruction, were compared at the three time points: Pre-, Post- and Delayed Post-Intervention, to see if any change had occurred Pre- to Post- intervention and whether it had been maintained three months after intervention in the Delayed Post-Intervention . It was hoped to ascertain whether either of the two conditions produced changes in pronunciation in the direction of nativeness according to the rater judgements. The data was analysed using descriptive statistics to identify any obvious trends and further analysed using inferential statistics (a mixed ANOVA) to identify whether any broader generalisations could be made about the data. It was predicted that top-down strategies to work on pronunciation would be more effective than bottom-up at modifying spontaneous speech in line with the recommendations of prominent researchers

(LaScotte et al., 2021; Moyer, 2018; Pennington, 2019) and the success of research into strategies to achieve pronunciation modification such as shadowing (Foote & McDonough, 2017; Martinsen et al, 2017), echoing (Chung, 2017), mirroring (Meyers, 2018), channeling (La Scotte & Tarone, 2019; Moreno, 2016) and roleplay (Pennington & Waxler, 2017).

5.3.1 Analysis of Overall Measure (All Six Measures Combined) for Group 1 and Group 2

A descriptive analysis of the data shows that, in terms of overall mean scores over the six measures (nativeness, segmental/word-level features, sentence-level features, fluency, vocabulary and accuracy), Group 2 was judged to be more advanced than Group 1 in the Pre-Intervention condition (Figure 10). Both groups experienced an increase Post-Intervention, but Group 1 showed the greatest overall increase. In the Delayed condition, Group 2 showed a slight drop in attainment, while Group 1 maintained earlier gains so that both groups were judged to be performing at a similar level in the Delayed Post-Intervention 3 months after the end of the intervention.

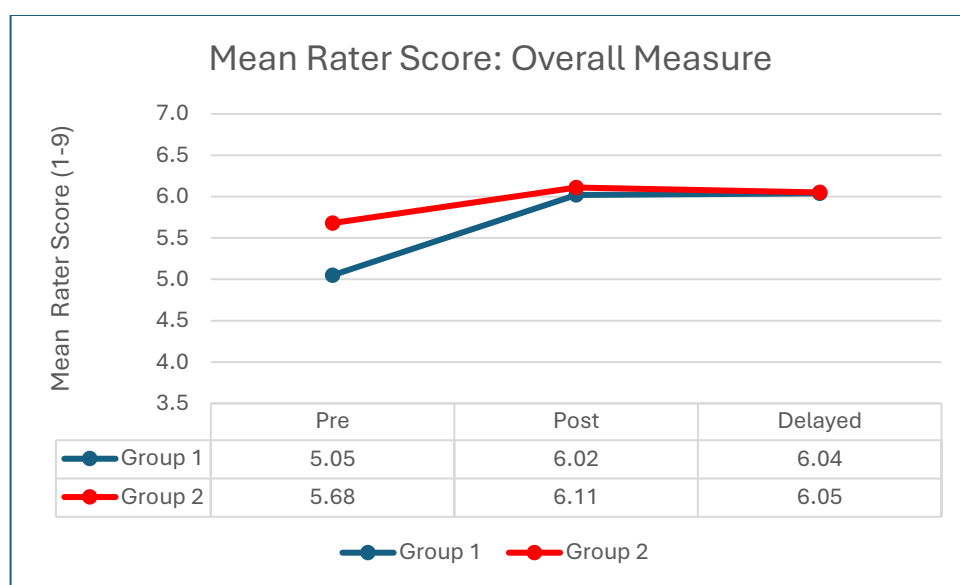


Figure 10. Mean Rater Score per Group: Overall Measures (rounded to 2 decimal places) - Figure 3 minus Control

The overall scores were further analysed using inferential statistics in the form of a 2x(3) mixed ANOVA, with the between-subjects factor of group (Group 1, Group 2) and the within-subjects factor of time (Pre-, Post-, Delayed-) in order to test if any general trends can be inferred from the data. The mixed ANOVA showed statistically significant main effects for time: $F(2,28)= 9.15$, $p < 0.001$, but not for group: $F(1,14)= 0.63$, $p = 0.44$ and there was no significant interaction between time and group $F(2, 28) = 1.62$, $p = 0.22$ (Table 30).

Table 30. Mixed ANOVA: Overall Measures – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	5.10	2	2.55	9.15	<0.001
Time x Group	0.90	2	0.45	1.62	0.22
Error (Time)	7.81	28	0.28		
Group	0.68	1	0.68	0.63	0.44
Error (Group)	15.05	14	1.07		

The results suggest that while both groups increased over time, there was no significant difference in attainment between the groups, so that one intervention was not superior to the other.

In order to give a more nuanced picture, in the sections which follow, each measure (nativeness, segmental/word-level features, sentence-level features, fluency, vocabulary, accuracy) is analysed separately.

5.3.2 Analysis of Measures Individually (Group 1 and 2)

5.3.2.1 Nativeness

Group 1 mean scores for nativeness were lower than Group 2 mean scores in the Pre-Intervention Condition showing a distance of 0.73 of a point on the 9-point rating scale (see Figure 11). This indicates that the raters initially viewed Group 2 as the stronger group in terms of nativeness. Both groups showed an increase in the Post-Intervention and these increases are maintained in the Delayed-Post Intervention condition.

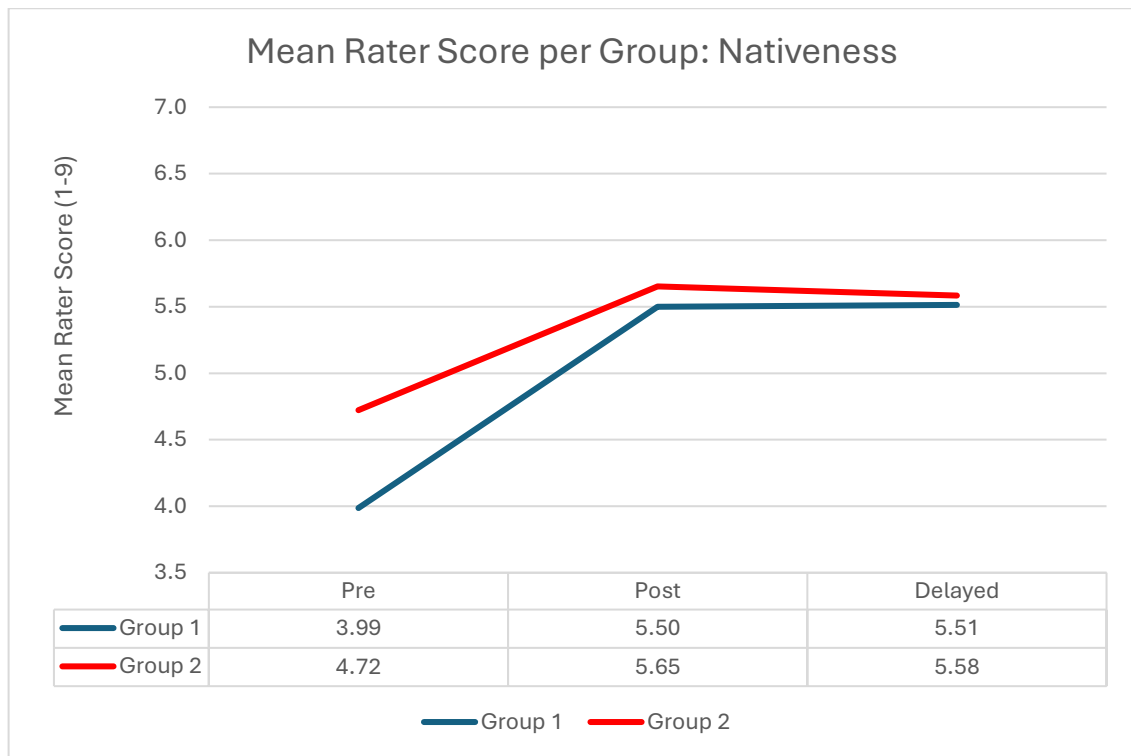


Figure 11. Mean Rater Score per Group: Nativeness (rounded to 2 decimal places) - Figure 4 minus Control

However, Group 1 showed a larger change from Pre to Post of 1.52 points as compared to 0.93 points for Group 2 (Table 31). Between Post and Delayed there was little overall change for either group. Group 1 showed a slight increase of 0.01, while Group 2 showed a slight drop of - 0.07. Overall, therefore Group 1 showed a larger overall rise in rater nativeness score from Pre to Delayed of 1.53 points compared to Group 2 who saw a rise of 0.86 of a point, suggesting that the Group 1 intervention was more effective in terms of nativeness rating.

Table 31. Mean change in Rater Score per Speaker: Nativeness - Groups 1 and 2 (rounded to 2 decimal places)

Group	Speaker	Mean Change in Rater Score: Pre to Post	Mean Change in Rater Score: Post to Delayed	Mean Change in Rater Score: Pre to Delayed
1	1	0.89	0.44	1.33
1	2	2.33	0.11	2.45
1	3	2.78	-0.44	2.33
1	4	1.00	-0.89	0.11
1	5	1.11	0.33	1.44
1	6	0.67	0.88	1.55
1	7	1.89	-0.89	1.00
1	8	1.45	0.55	2.00
1	MEAN	1.52	0.01	1.53
2	9	1.22	0.33	1.55
2	10	2.78	-1.22	1.56
2	11	1.11	0.44	1.55
2	12	0.45	0.11	0.56
2	13	1.00	-0.22	0.78
2	14	0.44	0.56	1.00
2	15	-0.11	-1.22	-1.33
2	16	0.56	0.67	1.23
2	MEAN	0.93	-0.07	0.86

Looking at individual differences in attainment for each participant (Table 31), it can be seen that every participant, except Speaker 15, increased from Pre to Post, which indicates that, while Group 1 made bigger nativeness gains, both interventions were effective.

Speakers 3, 4, 7, 10 and 13 saw a drop in rating between Post and Delayed, but in none of the cases did this result in a return to Pre-Intervention levels. The remaining speakers made gains from Pre to Post and gains from Post to Delayed, with the exception of Speaker 15. (See Section 5.2.4.1 for a discussion of this speaker.) Again, this is suggesting that, while there may be a slight drop in attainment in some cases in the Delayed condition, overall, the increases gained are maintained 3 months after the end of the intervention.

The mean nativeness scores were further analysed a 2 x (3) mixed ANOVA as before. There were statistically significant main effects for time $F(2, 28) = 24.22$, $p < 0.001$, but not for group $F(1,14) = 0.62$, $p = 0.44$. There was no significant interaction between time and group; $F(2, 28) = 1.64$, $p = 0.21$ (Table 32).

Table 32. Mixed ANOVA: Nativeness – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	15.58	2	7.79	24.22	<0.001
Time x Group	1.05	2	0.53	1.64	0.21
Error (Time)	9.00	28	0.32		
Group	1.22	1	1.22	0.62	0.44
Error (Group)	27.73	14	1.98		

Taken together, these results suggest that, while there was a statistically significant increase in both groups after the intervention, there was no significant difference between the groups. It would appear that, in terms of nativeness, both types of intervention were successful in achieving increases, but it could not be said that one approach was better than the other.

5.3.2.2 Segmental/Word-Level Features

In terms of the rating of the tokens for segmental/word-level features, as with nativeness, Group 1 scored lower overall than Group 2 at all three time-intervals (see Figure 12) suggesting that Group 2 was initially the better group and remained so throughout the treatment.

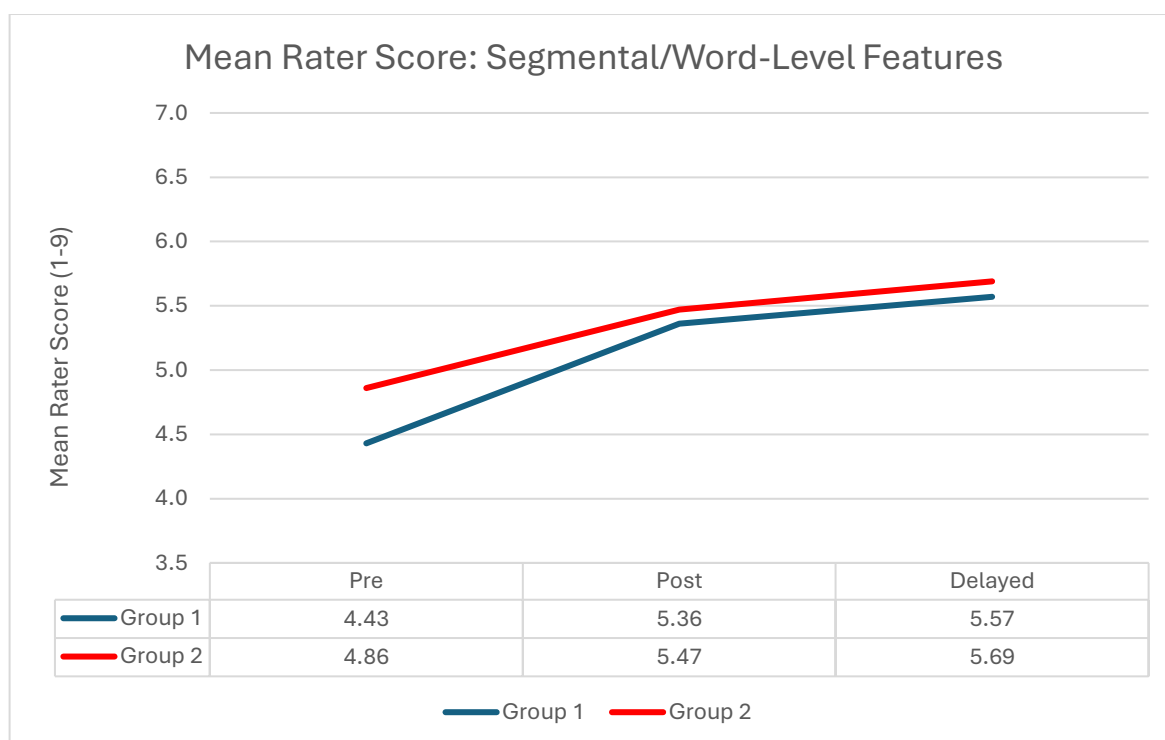


Figure 12. Mean Rater Score per Group: Segmental/Word-Level Features (rounded to 2 decimal places) - Figure 5 minus Control

The overall trend shows continued increase for both groups from Pre to Post and Post to Delayed. However, the bigger overall increase was achieved by Group 1, who increased by a

total of 1.12 points on the 1-9 rating scale from Pre to Delayed, whereas Group 2 only increased by 0.82 points (Table 33).

Table 33. Mean change in Rater Score per Speaker: Segmental/Word-Level Features – Groups 1 and 2 (rounded to 2 decimal places)

Group	Speaker	Mean Change in Rater Score: Pre to Post	Mean Change in Rater Score: Post to Delayed	Mean Change in Rater Score: Pre to Delayed
1	1	0.44	0.22	0.67
1	2	1.87	0.22	2.09
1	3	2.56	0.11	2.67
1	4	1.00	-1.11	-0.11
1	5	1.12	-0.11	1.01
1	6	-0.44	1.11	0.67
1	7	0.45	0.67	1.11
1	8	0.33	0.56	0.89
1	MEAN	0.92	0.21	1.13
2	9	0.67	1.22	1.89
2	10	2.22	-0.78	1.44
2	11	0.56	0.78	1.33
2	12	0.11	0.11	0.22
2	13	0.67	0.22	0.89
2	14	0.00	0.44	0.44
2	15	-0.44	-0.78	-1.22
2	16	1.00	0.56	1.56
2	MEAN	0.60	0.22	0.82

This trend was generally borne out by individual speakers in both groups with increases being made from Pre to Post and maintained or increased Post to Delayed. In Group 1, only Speaker 6 failed to show an increase from Pre to Post. However, they then showed an increase of 1.1 points from Post to Delayed giving an overall increase of 0.67 of a point. Speakers 4 and 5 increased from Pre to Post but showed a drop from Post to Delayed. The drop in rating for Speaker 5 was minimal (0.11), so the increase from Pre to Post was largely maintained with overall increase of 1.01 points. However, Speaker 4 showed a drop off from Post to Delayed of -1.11, so that the overall change from Pre to Delayed was a slight drop of -0.11 of a point on the 1-9 Rating Scale. Speaker 4 was, therefore, the only participant in Group 1 to fail to make any increase in this area.

In Group 2, with the exception of Speaker 15 who showed a gradual decline over the 3 time periods (see discussion of Speaker 15 in 5.2.4.1 above), every speaker showed an increase between Pre and Post, except Speaker 14 who showed no change. This increase was maintained

in the Delayed Post Test, except by Speaker 10 who showed a drop of -0.78 of a point from Post to Delayed, but still showed an overall increase of 1.44 points from Pre to Delayed.

Taken as a whole, therefore, both groups showed an increase from Pre to Post and maintained or increased the increase from Post to Delayed, but there were outliers. The only speaker to show a marked drop from Pre to Delayed was Speaker 15.

The mean scores for segmental/word-level features were analysed using a 2 x (3) mixed ANOVA. There were statistically significant main effects for time $F(2, 28) = 11.84$, $p < 0.001$, but not for group $F(1, 14) = 0.36$, $p = 0.56$. There was no significant interaction between time and group; $F(2, 28) = 0.37$, $p = 0.69$ (Table 34).

Table 34. Mixed ANOVA: Segmental/Word-Level Features – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	8.34	2	4.17	11.84	<0.001
Time x Group	0.26	2	0.13	0.37	0.69
Error (Time)	9.86	28	0.35		
Group	0.59	1	0.59	0.36	0.56
Error (Group)	22.95	14	1.64		

Therefore, as for nativeness, both groups showed significant increase over time, but there was no significant difference between the two interventions on outcome.

5.3.2.3 Sentence-Level Features

Sentence-level features on average showed an increase from Pre to Delayed, with Group 1 showing the greater overall increase (Figure 13).

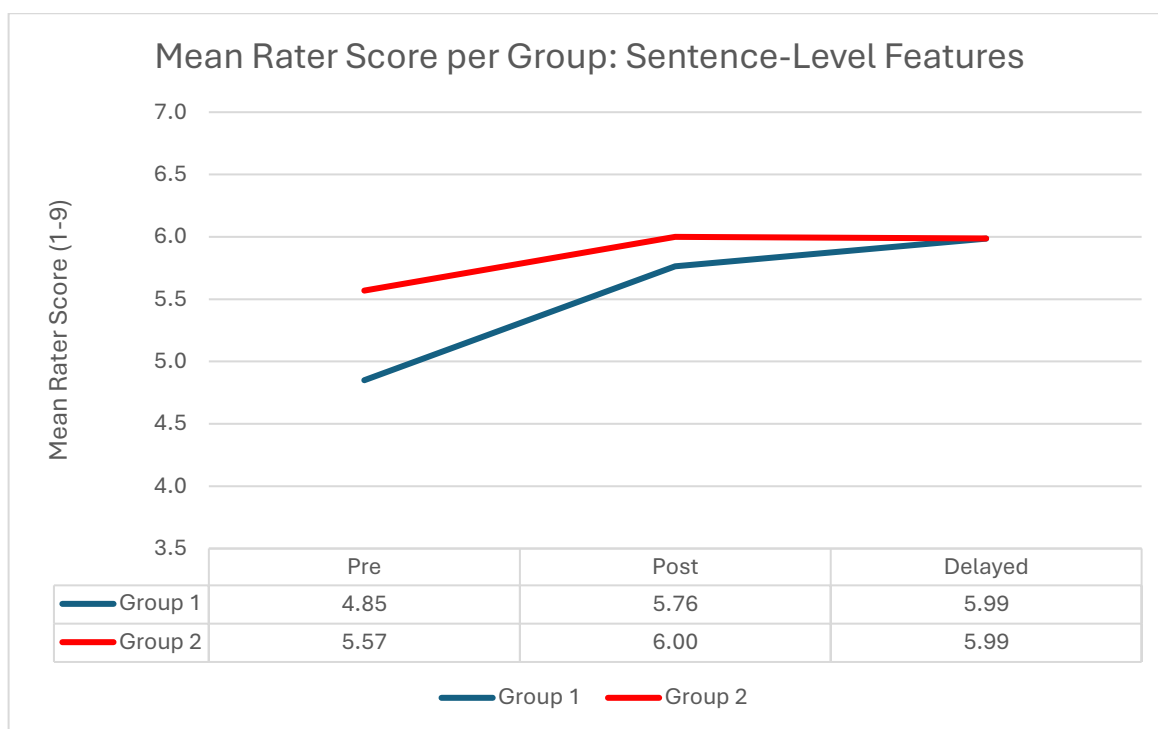


Figure 13. Mean Rater Scores per Group: Sentence-Level Features (rounded to 2 decimal places) - Figure 6 minus Control

In terms of individual differences, all speakers in Group 1 showed an increase from Pre to Post except Speakers 3 and 4 who showed a drop in attainment (Table 35). Only Speaker 4 maintained that decline from Pre to Delayed showing a slight decline of -0.33 of a point. In Group 2, several speakers showed a drop or no change in scores, but again, it is only Speaker 15 who shows an overall loss from Pre to Delayed (see Section 5.2.4.1 for a discussion of Speaker 15). The overall trend, therefore, is towards gains being made in both groups, but with Group 1 showing the greater gain.

Table 35. Mean change in Rater Score per Speaker: Sentence-Level Features – Groups 1 and 2 (rounded to 2 decimal places)

Group	Speaker	Mean Change in Rater Score: Pre to Post	Mean Change in Rater Score: Post to Delayed	Mean Change in Rater Score: Pre to Delayed
1	1	0.45	0	0.45
1	2	1.00	0.56	1.56
1	3	2.33	-0.44	1.89
1	4	0.00	-0.33	-0.33
1	5	0.67	0	0.67
1	6	0.78	1.00	1.78
1	7	0.56	0.22	0.78
1	8	1.56	0.78	2.33
1	MEAN	0.92	0.22	1.14
2	9	1.22	0.44	1.67
2	10	1.33	-0.67	0.67
2	11	-0.22	0.33	0.11
2	12	0.78	-0.67	0.11
2	13	0.45	0	0.45
2	14	0.11	0.78	0.89
2	15	-0.67	-0.55	-1.22
2	16	0.45	0.22	0.67
2	MEAN	0.43	-0.01	0.42

The mean sentence-level scores were analysed using a 2 x (3) mixed ANOVA. There were statistically significant main effects for time $F(2, 28) = 1.21$, $p < 0.001$, but not for group $F(1,14) = 1.10$, $p = 0.31$. There was no significant interaction between time and group; $F(2, 28) = 2.13$, $p = 0.14$ (Table 36).

Table 36. Mixed ANOVA: Sentence-Level Features – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	5.71	2	2.85	1.21	<0.001
Time x Group	1.08	2	0.54	2.13	0.14
Error (Time)	7.13	28	0.25		
Group	1.23	1	1.23	1.10	0.31
Error (Group)	15.58	14	1.11		

Therefore, the inferential results suggest that both groups achieved higher mean rater scores over time, but there is no significant difference between the two interventions and no effect of the different interventions on outcome.

5.3.2.4 Fluency

Fluency ratings were markedly different between the two groups. While Group 1 showed a large increase in fluency from Pre to Post (1.24 points), which was largely maintained in the Delayed Post-Test, Group 2 scores for fluency remained largely unchanged at the three time points (Figure 14).

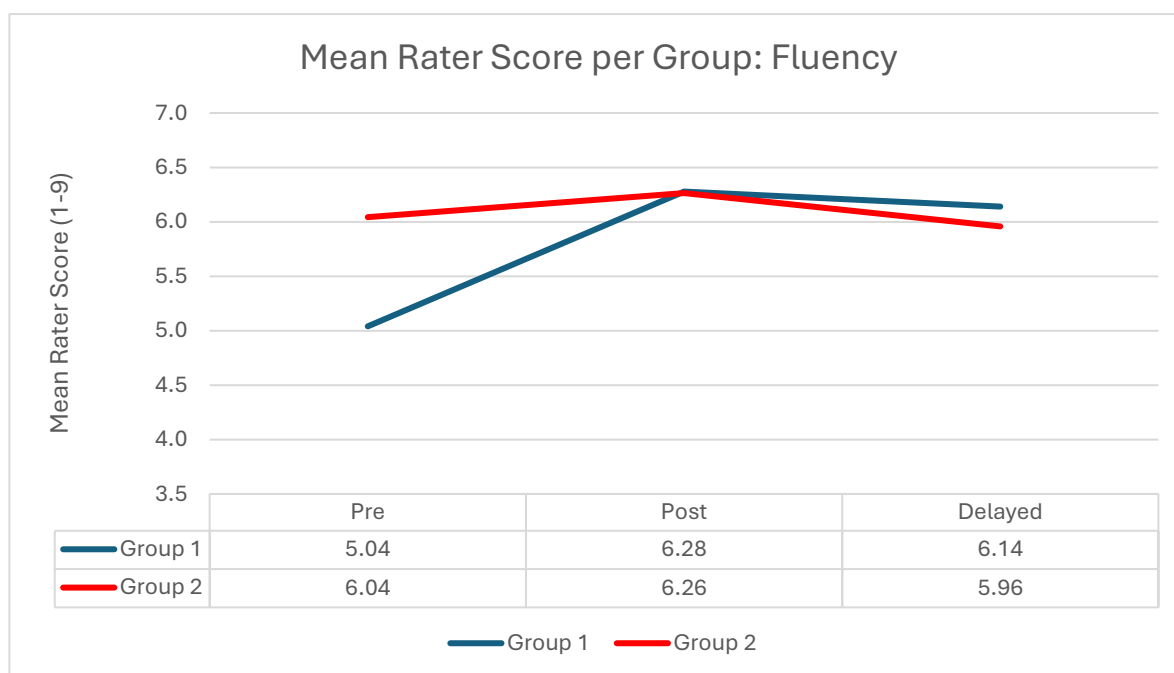


Figure 14. Mean Rater Score per Group: Fluency (rounded to 2 decimal places) - Figure 7 minus Control

In Group 1, all speakers except Speaker 4 made increases from Pre- to Delayed Post-Test, with Speaker 3 improving by 2.56 points and Speaker 8 improving by 3 points (Table 37). In Group 2, changes were largely small with the exception of Speaker 9, who increased by 1.78 points, and Speaker 15, who dropped by 2.11 points (see Section 5.2.4.1 for a discussion of Speaker 15). Group 1, therefore, seems to have benefitted more from the intervention than Group 2 in terms of fluency.

Table 37. Mean Change in Rater Score per Speaker: Fluency – Groups 1 and 2 (rounded to 2 decimal places)

Group	Speaker	Mean Change in Rater Score: Pre to Post	Mean Change in Rater Score: Post to Delayed	Mean Change in Rater Score: Pre to Delayed
1	1	1.11	-0.78	0.34
1	2	1.11	0.33	1.44
1	3	3.12	-0.56	2.56
1	4	-0.22	-0.67	-0.88
1	5	1.67	-1.00	0.67
1	6	0.82	0.45	1.26
1	7	0.11	0.33	0.45
1	8	2.22	0.78	3.00
1	MEAN	1.24	-0.14	1.10
2	9	0.67	1.11	1.78
2	10	2.56	-2.00	0.56
2	11	-0.78	0.22	-0.56
2	12	0.56	-0.44	0.11
2	13	0.78	-0.67	0.11
2	14	-0.78	0.56	-0.22
2	15	-1.44	-0.67	-2.11
2	16	0.22	-0.56	-0.33
2	MEAN	0.22	-0.31	-0.08

The mean fluency scores were analysed using a 2 x (3) mixed ANOVA. There were statistically significant main effects for time $F(2, 28) = 4.09$, $p < 0.028$, and for group $F(1, 14) = 0.59$, $p = 0.041$. There was no significant interaction between time and group; $F(2, 28) = 3.01$, $p = 0.065$ (Table 38).

Table 38. Mixed ANOVA: Fluency – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	4.66	2	2.33	4.09	0.028
Time x Group	3.44	2	1.72	3.01	0.065
Error (Time)	15.98	28	0.57		
Group	0.93	1	0.93	0.59	0.041
Error (Group)	22.02	14	1.57		

Therefore, the ANOVA identifies a significant difference between the two groups in terms of fluency. At first glance, it would appear that the bottom-up intervention received by Group 1 was more effective at improving fluency than the top-down intervention received by Group 2

and this finding is statistically significant. However, the result has to be looked at with caution: fluency was not explicitly targeted by the intervention, so this seems to be a strange result. A possible explanation is that, as stated in Section 5.2.5, Group 1 were the older group having a mean age of 31.1 years at the start of the study, while Group 2 had a mean age of 24.8 years. In Section 5.2.5 it was proposed that, as younger Portuguese learners of English are more exposed to NS English than was the case in the past, this could have meant that Group 2 were initially the more fluent group. This being the case, Group 1 may have benefitted more from interaction with NS English, compared to Group 2, who were already judged as quite fluent prior to intervention.

Nevertheless, fluency was the only measure where a statistically significant interaction was found for time and group, meaning that, in contrast with results obtained for other measures, the group a participant was in influenced attainment.

5.3.2.5 Vocabulary

For vocabulary, Group 2 appeared to remain relatively unchanged by the intervention, although Group 1 showed a 0.65-point rise from Pre- to Post-Test which was maintained in the Delayed Post-Test (Figure 15).

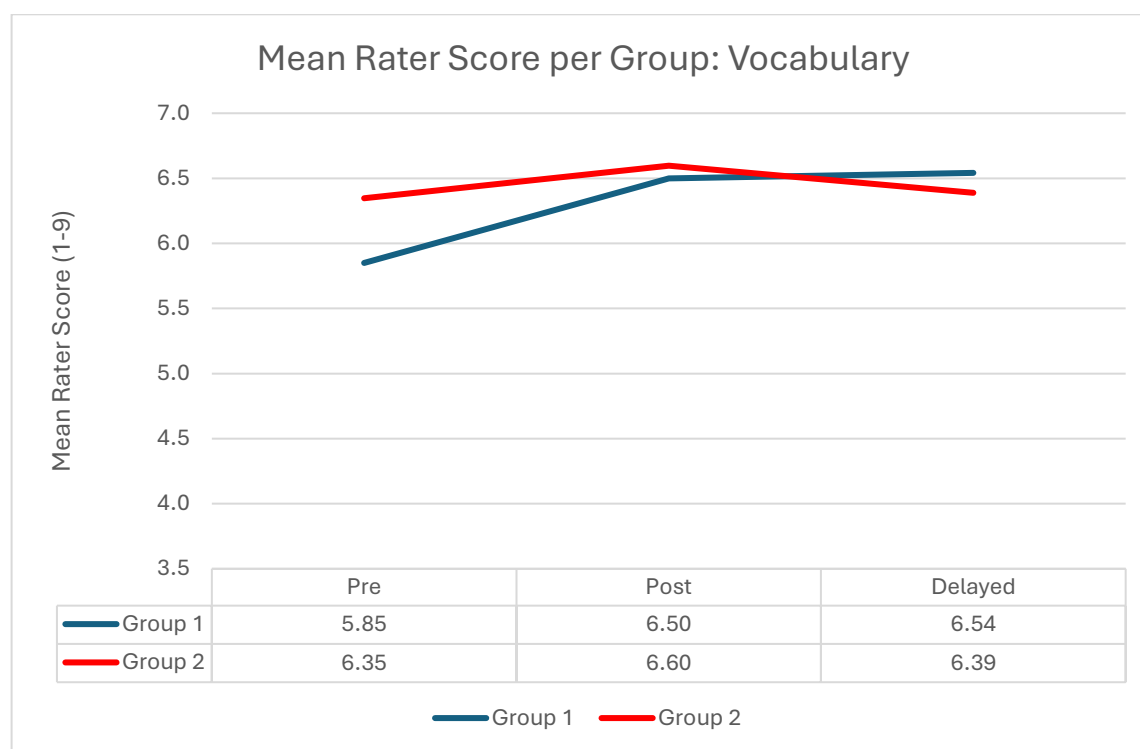


Figure 15. Mean Rater Score per Group: Vocabulary (rounded to 2 decimal places) - Figure 8 minus Control

There was some fluctuation in individual scores (Table 39). Speakers 3,6,8 and 9 achieved gains of over one point, while Speaker 15 dropped by 1.78 points from Pre- to Delayed Post-Test (see Section 5.2.4.1 for a discussion of Speaker 15). Taken as a whole, however, vocabulary scores showed little change.

Table 39. Mean Change in Rater Score per Speaker: Vocabulary – Groups 1 and 2 (rounded to 2 decimal places)

Group	Speaker	Mean Change in Rater Score: Pre to Post	Mean Change in Rater Score: Post to Delayed	Mean Change in Rater Score: Pre to Delayed
1	1	0.33	0.44	0.78
1	2	0.89	0.00	0.89
1	3	2.00	-0.56	1.44
1	4	0.33	-0.89	-0.56
1	5	0.22	-1.00	-0.78
1	6	0.67	0.78	1.44
1	7	0.11	0.00	0.11
1	8	0.67	1.56	2.22
1	MEAN	0.65	0.04	0.69
2	9	1.11	0.33	1.45
2	10	2.00	-1.56	0.44
2	11	0.11	0.11	0.22
2	12	0.56	0.00	0.56
2	13	0.67	-0.56	0.11
2	14	-0.33	-0.11	-0.44
2	15	-1.89	0.11	-1.78
2	16	-0.22	0	-0.22
2	MEAN	0.25	-0.21	0.04

This was also the finding of a 2 x (3) mixed ANOVA. There were no statistically significant results for time $F(2, 28) = 2.33$, $p = .116$, or group $F(1,14) = .36$, $p = .55$. There was no significant interaction between time and group; $F(2, 28) = 1.1$, $p = .348$ (Table 40).

Table 40. Mixed ANOVA: Vocabulary – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	1.85	2	0.92	2.33	0.116
Time x Group	0.87	2	0.43	1.1	0.348
Error (Time)	11.08	28	0.4		
Group	0.26	1	0.26	0.36	0.55
Error (Group)	10.07	14	0.72		

This is an important result. The intervention did not attempt to work on vocabulary enrichment in either group, so the lack of a significant result for Time, Group or the Time x Group interaction is to be expected.

5.3.2.6 Accuracy

Mean rater scores for accuracy remained largely unchanged over the three time points and were similar for each group, hovering above 6 points for both groups (Figure 16).

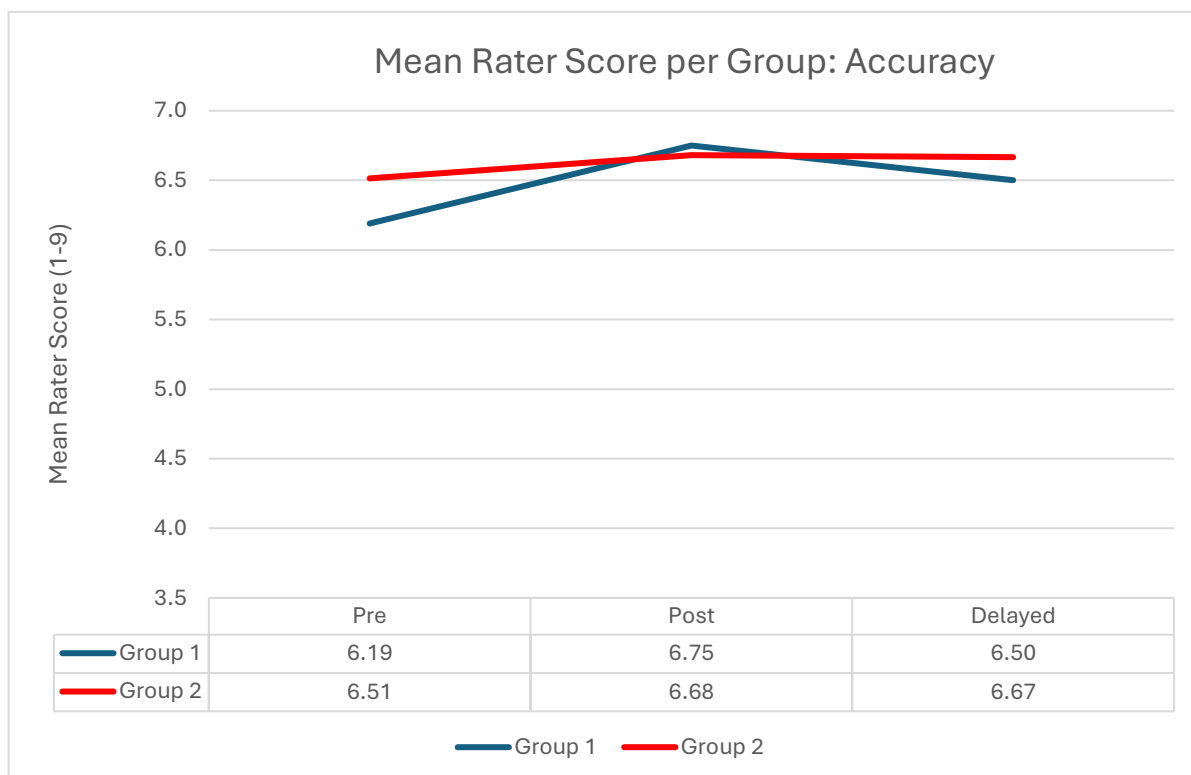


Figure 16. Mean Rater Score per Group: Accuracy (rounded to 2 decimal places) - Figure 9 minus Control

There were individual differences between speakers with Speaker 3, 8 and 12 making gains of over 1 point on the rating scale from Pre to Delayed condition and Speaker 15 suffering a loss of over 1 point (see Section 5.2.4.1 for a discussion of Speaker 15). However, the difference in mean scores for accuracy over the three time points was minimal (Table 41).

Table 41. Mean Change in Rater Score per Speaker: Accuracy – Groups 1 and 2 (rounded to 2 decimal places)

Group	Speaker	Mean Change in Rater Score: Pre to Post	Mean Change in Rater Score: Post to Delayed	Mean Change in Rater Score: Pre to Delayed
1	1	0.89	-0.22	0.67
1	2	0.45	-0.56	-0.11
1	3	1.56	-0.44	1.11
1	4	0.44	-0.67	-0.22
1	5	0.67	-1.33	-0.67
1	6	-0.22	0.78	0.56
1	7	0.11	-0.22	-0.11
1	8	0.56	0.67	1.22
1	MEAN	0.56	-0.25	0.31
2	9	0.78	0.11	0.89
2	10	1.67	-1.33	0.33
2	11	-0.22	0.00	-0.22
2	12	1.00	0.11	1.11
2	13	0.33	-0.22	0.11
2	14	-1.11	0.67	-0.44
2	15	-1.22	0	-1.22
2	16	0.11	0.56	0.67
2	MEAN	0.17	-0.01	0.15

As previously, the mean accuracy scores were analysed using a 2 x (3) mixed ANOVA. There were no statistically significant main effects for time $F(2, 28) = 2$, $p=0.154$, or group $F(1,14) = 0.28$, $p = 0.604$. There was no significant interaction between time and group; $F(2, 28) = 0.58$, $p = 0.568$ (Table 42).

Table 42. Mixed ANOVA: Accuracy – Groups 1 and 2 (rounded to 2 decimal places)

	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F ratio (F)	p-value (p)
Time	1.07	2	0.53	2	0.154
Time x Group	0.31	2	0.15	0.58	0.568
Error (Time)	7.47	28	0.27		
Group	0.23	1	.23	0.28	0.604
Error (Group)	11.52	14	0.82		

As for vocabulary, the accuracy results show a lack of significant findings for Time, Group or Time x Group interaction. The intervention did not attempt to work on language accuracy in either group, so the lack of a significant result for Time, Group or the Time x Group interaction is to be expected.

5.3.4 Discussion – Question 3

The results indicate a positive picture in terms of intervention, in that both courses resulted in statistically significant gains in the targeted areas of nativeness, individual segmental/word-level features and sentence-level features, over the three time periods. This suggests that both interventions were effective in addressing pronunciation.

The research did not, however, show that one group was significantly more effective than the other for the three measures over the three time periods. While, on the whole, Group 1 made the greater gains, these were shown not to be statistically significant under inferential analysis with the exception of the fluency measure. However, as discussed in 5.2.5, as fluency was not a target of instruction, age differences in the groups would appear to be the most logical explanation for this outlier.

The failure to find a distinction in results between the two groups suggests that both approaches were effective in achieving accent modification.

5.3.5 Conclusion – Question 3

In answer to Research Question 3: *Will a top-down strategy-led approach bring late learners closer to achieving a native-like accent than a bottom-up structure-led approach?*, it was predicted that the top-down approach would prove superior in line with the recommendations of prominent researchers in the field of pronunciation (LaScotte et al., 2021; Moyer, 2018; Pennington, 2019). However, the prediction was not confirmed by the research. While there were individual differences in the raters' judgements of the measures analysed, both intervention programmes were effective in achieving a statistically significant gains in nativeness, which seemed to be a result of changes in segmental/word-level features and sentence-level features. The group difference obtained for 'fluency' was judged to be most likely a result of group mean age differences prior to the course. Vocabulary and accuracy appeared to be largely unaffected by the intervention.

5.4 QUESTION 4: CAN INTERVENTION INFLUENCE THE LEARNERS' ABILITY TO EVALUATE THEIR OWN SPONTANEOUS SPEECH PRODUCTION TO CORRESPOND TO THAT OF NATIVE SPEAKER RATERS?

To answer this question, the participants in the two intervention groups completed an End of Course Feedback Questionnaire (Appendix E), which asked them to comment on their awareness of their own accent (Question 1), which areas of pronunciation the intervention helped them to work on (Question

2), what accent they would like to have in English (Question 3), if they felt the course helped them to modify their pronunciation (Question 4), what aspects of their pronunciation it had helped them change (Question 5) and which activities they found most useful in working on pronunciation (Question 6) along with space for extra comments. It was predicted that the intervention which specifically encouraged participants to listen to their Pre-Intervention token (Lesson 1), gave teacher and peer feedback on activities during the course and which gave teacher, rater and peer feedback on the Pre-Intervention token (Lesson 10) would help give the participants increased self-awareness of how others were judging their accent in line with findings of other researchers (Couper, 2003; Lappin-Fortin & Rye, 2014; MacDonald, 2018; McGregor & Reed, *ibid.*; Sardegna & McGregor, 2013; Tsunemoto et al., 2022).

5.4.1 End of Course Feedback Questionnaire

The End of Course Feedback Questionnaire (Appendix E) provided interesting insights into participants' impressions of the course and, more specifically their opinions on changes to their phonological self-awareness. There follows a discussion of the feedback for each of the six questions.

5.4.1.1 Question 1: *Do you feel the course has given you an increased awareness of your accent in English?*

All sixteen speakers answered 'yes' to this question, so, the participants of both groups felt that the intervention was effective in raising awareness to their current accent.

5.4.1.2 Question 2: *Which features of your accent do you have an increased awareness of?*

In Group 1 (Table 43), all participants stated that they had an increased awareness of individual sounds and sentence-level features (Speaker 7 did not tick the box for sentence-level features, but did specify coarticulation, which is a suprasegmental feature). Interestingly, only three speakers mentioned word stress.

Table 43. Group 1: Which features of your accent do you have an increased awareness of?

Group 1	Individual Sounds	Word stress	Sentence-level	Fluency	Other
Speaker 1	✓		✓		
Speaker 2	✓	✓	✓	✓	
Speaker 3	✓		✓		
Speaker 4	✓	✓	✓	✓	
Speaker 5	✓		✓		
Speaker 6	✓		✓	✓	
Speaker 7	✓		(✓)	✓	Coarticulation
Speaker 8	✓	✓	✓		

The picture for Group 2 is more erratic (Table 44). While all speakers bar Speaker 13 felt they had an awareness of segments, only six mentioned sentence-level features, five ticked word stress and four mentioned fluency. This was possibly because the whole focus of the course was less on specific features of production and more on working on production as a whole.

Table 44. Group 2: Which features of your accent do you have an increased awareness of?

Group 2	Individual Sounds	Word stress	Sentence-level	Fluency	Other
Speaker 9	✓	✓	✓	✓	
Speaker 10	✓	✓	✓		
Speaker 11	✓	✓	✓		
Speaker 12	✓	✓	✓	✓	
Speaker 13			✓	✓	
Speaker 14	✓		✓		
Speaker 15	✓				
Speaker 16	✓	✓		✓	

5.4.1.3 Question 3: *What would be the accent which you would like to have in English?*

Interestingly, while at the start of the course, all the participants stated British English as their goal as discussed in Section 4.3, at the end of the intervention, only 10 of the 16 still maintained this goal, five in Group 1, and five in Group 2, suggesting there was no difference between the two groups regarding this finding. Three participants stated they were happy with their current accent, Speaker 1 stated Standard British or American as the goal, Speaker 11 stated no particular accent and Speaker 16 had changed to American English. It could be that, by attempting to give participants greater pronunciation self-awareness, they decided that they were satisfied with their current accent. Alternatively, it is possible that the interventions

showed the participants that pronunciation modification requires a considerable amount of practice which some may have been unwilling or unable to dedicate time to.

5.4.1.4 Question 4: *Do you feel that the course has helped you to change your pronunciation?*

13 of the 16 participants replied ‘yes’ to this question. Speaker 1 stated that while the course had increased awareness, they were not sure if the accent had changed. Speaker 5 replied that they were happy with their current accent. Speaker 16 ticked both ‘yes’ and ‘no’, but this may have been a mistake since the participant then ticked word stress and fluency as having changed.

5.4.1.5 Question 5: *If the course helped you to change, what aspects of your pronunciation has it helped you to change?*

Answers for Group 1 were heavily biased towards individual sounds, so all speakers, except Speaker 5, who had decided they were happy with their accent, felt they had achieved changes at a segmental level, as shown in Table 45. There were only a few mentions of suprasegmentals being influenced by the instruction, two for word stress, two for sentence-level features and three for fluency.

Table 45. Group 1: If the course helped you to change, what aspects of your pronunciation has it helped you to change?

Group 1	Individual Sounds	Word Stress	Sentence-Level	Fluency	Other
Speaker 1	✓				
Speaker 2	✓	✓			
Speaker 3	✓		✓		
Speaker 4	✓			✓	
Speaker 5					
Speaker 6	✓			✓	
Speaker 7	✓				Coarticulation
Speaker 8	✓	✓	✓	✓	

These findings for Group 1 were compared to participant responses to Pre-Intervention Pronunciation Questionnaire Question 8 (Table 46) which asked *do you have difficulty with any particular sounds or features of spoken English?* The findings show that five of the participants felt they had problems at the level of individual sounds prior to intervention, but all felt they had achieved modification in this measure, with the exception of Speaker 5 who, as previously stated, stated they were now happy with their current accent. This suggests that two

speakers (Speakers 3 and 6) who had not previously seen this as a problem area, now recognised it as such and had made changes, while the other five (Speakers 1, 2, 4, 7 and 8) saw this as a problem prior to intervention and felt the intervention had helped them to change. This is probably a reflection of the fact that four Lessons (2, 3, 6 and 9) looked at segmentals. However, the picture for the other measures is less clear. For Word Stress, Speakers 1 and 2 highlighted this as a problem area, but only Speaker 2 felt the intervention had provoked change in this area. Speaker 8, who did not see this as a difficulty prior to intervention, felt that the course had helped with modification. The findings for Word Stress may be a reflection of the fact that, as mentioned before, it was only specifically addressed in Lesson 4. For sentence-level features Speakers 2, 3, 6 and 7 saw this as a problem prior to intervention, but of them, only Speakers 3 and 7 felt the course helped them to change in this area. Speaker 8, who had not identified sentence-level features as a problem prior to the intervention, nevertheless stated they felt the intervention had led to change in this area. This is a somewhat surprising finding given that, as mentioned in Section 5.4.2.2, four lessons (4, 5, 7, 8) looked at this feature. It is possible that the structural approach to the teaching of suprasegmentals, which followed an FonFS approach was not felt to be effective by the majority of participants. Finally, fluency, which was identified as a problem prior to intervention by Speakers 2, 4 and 5 was felt to have been modified by the intervention by Speakers 4, 6 and 8. Since fluency was not specifically addressed in the intervention, little change was expected, so this finding for Speakers 4, 6 and 8 is possibly a reflection of the fact that the intervention involved a considerable amount of spontaneous interaction, as discussed in Section 5.2.5 above, which may have inadvertently helped some students in this area.

Table 46. Group 1 Comparison of Responses between Pre-Intervention Questionnaire Question 8 and Post-Intervention End of Course Feedback Question 5

Group 1	Individual Sounds		Word Stress		Sentence Level		Fluency		Other	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Speaker 1	✓	✓	✓							
Speaker 2	✓	✓	✓	✓	✓		✓			
Speaker 3		✓			✓	✓				
Speaker 4	✓	✓					✓	✓		
Speaker 5							✓			
Speaker 6		✓			✓			✓		
Speaker 7	✓	✓			✓	(✓)				Co-articulation
Speaker 8	✓	✓		✓		✓		✓		

For Group 2, the picture was less homogeneous with six speakers mentioning change at a segmental level, but also two felt they had achieved change in word stress, five at the sentence level and three in fluency (see Table 47). Overall, therefore, while Group 1 and Group 2 are comparable in terms of segmental change, Group 2 were superior in the changes they felt they had made at the suprasegmental level.

Table 47. Group 2: If the course helped you to change, what aspects of your pronunciation has it helped you to change?

Group 2	Individual Sounds	Word Stress	Sentence-Level	Fluency	Other
Speaker 9	✓	✓	✓	✓	
Speaker 10	✓		✓		
Speaker 11	✓				
Speaker 12	✓		✓		
Speaker 13			✓	✓	
Speaker 14	✓		✓		
Speaker 15	✓				
Speaker 16		✓		✓	

These findings were compared to participant responses to Pre-Intervention Questionnaire Question 8 (Table 48) which asked *do you have difficulty with any particular sounds or features of spoken English?* There appeared to be a close correlation between the areas highlighted as problematic prior to intervention and the areas the speakers felt they had modified as a result of the intervention. At the level of individual sounds, all five speakers who highlighted this as a problem area reported change in this area. Equally, for word stress, Speaker 9 and Speaker 16 identified this as a problem prior to intervention and felt it had changed as a result of the intervention. Only Speaker 10 reported change in this area having not identified it as a problem prior to intervention. Speakers 10, 12 and 13 identified sentence-level features as problematic prior to intervention and felt it had been changed by the intervention, while Speakers 9 and 14 reported change in this area, having not identified it as a problem area prior to intervention. Fluency was felt to have changed for three speakers (9, 13 and 16) but only Speaker 9 had identified this as a problem prior to intervention. Overall, therefore, for Group 2, there appeared to be quite a close correlation between the features of pronunciation which participants had identified as problematic prior to intervention and the ones they felt had improved as a result of the intervention.

Table 48. Group 2 Comparison of Responses between Pre-Intervention Questionnaire Question 8 and Post-Intervention End of Course Feedback Question 5

Group 2	Individual Sounds		Word Stress		Sentence-Level		Fluency		Other	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Speaker 9	✓	✓	✓	✓		✓	✓	✓		
Speaker 10	✓	✓		✓	✓	✓				
Speaker 11	✓	✓								
Speaker 12	✓	✓			✓	✓				
Speaker 13					✓	✓		✓		
Speaker 14	✓	✓				✓				
Speaker 15										
Speaker 16			✓	✓				✓		

5.4.1.6 Question 6: *Which activities were most useful in helping you to work on pronunciation?*

Possibly the most important finding of the whole research, the feedback shows a patchwork of answers, with different participants highlighting different approaches which they found useful. Table 49 shows the replies for Group 1 and Group 2. While identifying an approach as useful is not the same as actually showing empirically that it is effective in achieving accent change, it is nonetheless important. An intervention which the learners do not believe in and are not motivated by will be unlikely to be successful. (See Section 2.7.1 for a discussion of the role of motivation in accent modification.) What this finding indicates is that different approaches seem to ‘speak’ to different learners. With the possible exception of looking at factors which influence L2 accent (which was only carried out with Group 2), all the approaches were found to be useful by at least two students in each group, with evaluating current accent being almost unanimously judged as useful.

Table 49. Group 1 and 2: Which activities were most useful in helping you to work on pronunciation?

Group	Speaker	Listen to target	Segmental	Supra-segmental	Compare BrE/American	Evaluate current accent	Factors influencing L2 accent	Holistic strategies
1	1	✓			✓			
	2	✓	✓		✓	✓		
	3	✓	✓	✓		✓		✓
	4	✓	✓			✓		
	5	✓	✓			✓		
	6	✓						
	7					✓		✓
	8		✓	✓	✓	✓		
2	9			✓				
	10	✓	✓	✓				✓
	11		✓			✓		✓
	12	✓		✓	✓	✓		✓
	13	✓		✓		✓	✓	
	14		✓	✓		✓		✓
	15		✓			✓		
	16	✓			✓	✓		✓

That looking at factors influencing accent was not mentioned by anyone in Group 1 is not surprising, since this was not mentioned in the intervention, but in Group 2, it was only judged useful by one participant (Speaker 13). In Group 2, factors influencing accent were mentioned in several lessons (Lesson 5: L1 Interference, Lesson 8: Accent and Identity) but were explicitly addressed in Lesson 6. Possibly, this lesson went into too much theoretical detail and materials covered could have been integrated into other, more practical lessons. For example, social identity could have been looked at in Lesson 5 which focused on L1 interference and motivation could have been looked at in Lesson 8 on accent and identity.

However, overall, the findings suggest that there is no such thing as a ‘silver bullet’ when it comes to a complex skill such as L2 accent acquisition. It is trite but true to say that we are all different and therefore it seems reasonable to assume that, while one approach benefits one learner, it may be ineffective with another. One wonders, for example, whether Speaker 15, who raters identified as using specific non-standard phonemes, would have benefitted more from a bottom-up approach focussing on segmentals/suprasegmentals, rather than the more strategy-based, top-down instruction of Group 2.

5.4.1.7 Question 6: Other Comments

At the end of Question 6, there was space allowed for the participants to “*add any other comments you have about the course*”. Seven of the participants made no comment at this stage. Five participants left comments giving thanks for the course as follows:

Speaker 4: *It was a really great experience to improve the accent and, to be honest, I wouldn't really change a thing, it was perfect and really helpful.*

Speaker 5: *I really enjoyed it. I started out wanting to change my accent but now feel much happier with it and I would like to only work on some individual sounds.*

Speaker 6: *Thank you so much for the opportunity! This course has helped me in many ways.*

Speaker 7: *It was very interesting and helpful all that we have learned. I will benefit greatly from this course if I become a [sic] English teacher in the future.*

Speaker 8: *It was important to learn that pronunciation is a skill.*

Speaker 16: *I loved the course. It helped me realise many things about BrE which I was not aware of. I will continue to try to be close to BrE accent by use [sic] some methods we learned during the course.*

The other three participants left suggestions, which will be looked at individually. Participant 1 (Group 1) said “*in general, there was not much time dedicated to each of the activities, and this might be a factor that precluded the observation of effects*”. This was also my own feeling regarding the course. 10 hours of intervention was relatively short and resulted in limited time for individual practice during lesson time. This issue will be discussed further in Section 5.5 below. Speaker 11 (Group 2) said the following: “*sadly I was very busy during the time the course took place, so I could no [sic] spend on it the time it deserved. Still, it is an experience I don't regret. I just wish I had more availability*”. Clearly, this is a problem with any teaching intervention if participants do not have the time and energy to attend classes, participate in activities and complete homework tasks. If, as has been argued, a large component of pronunciation modification is essentially skill acquisition (Section 3.1), accent is not something in which change would be expected overnight and involves practice on the part of the learner. This issue is discussed further in Section 5.5. Finally, Speaker 13 (Group 2) made the following comment: “*The balance of theory and practice was great. Listening to out [sic] own recording was very important in facilitating changes in our accents. If anything, there could have been a moment, right in the first lesson, where our recordings could be discussed together, rather than leaving it*

for the last lesson. In any case, the importance of listening to our own voices was indeed stressed throughout. The neutral, judgement-free stance on accents was well conveyed and did not detract from the overall message that speakers are capable of changing their accents, should they be willing to put some effort into it. All in all, a fantastic and very instructive course!” The participant made the valid point that the Pre-Intervention Token recordings could have been the focus of Lesson 1, rather than Lesson 10. This was, indeed, the intention but unfortunately, some of the raters were very slow in giving their feedback, so the rater feedback had to be delayed until the final lesson. It was felt that listening to the tokens in Lesson 1 for teacher feedback and then again in Lesson 10 for the rater feedback would feel like a repetition to the participants. Instead, participants were asked to introduce themselves and were given peer and teacher feedback on their pronunciation in response to that. This was not really satisfactory for two reasons. Firstly, the participants had only just met each other and, possibly because of this, appeared reluctant to ‘criticise’ each other. Feedback at this stage was therefore largely limited to teacher feedback. Secondly, the introduction given by the participants was not recorded and it was therefore not possible to allow participants to re-listen to their output. It would be advisable to take steps to ensure rater feedback on the Pre-Intervention token is received prior to commencement of the intervention.

5.4.2 Discussion – Question 4

In answer to Question 1 *do you feel the course has given you an increased awareness of your accent in English*, all sixteen speakers answered ‘yes’ to this question, so the participants of both groups felt that the intervention was effective in raising awareness to their current accent. When asked to specify which features of accent they had an increased awareness of (Question 2), all participants of Group 1 stated that they had a greater awareness of segmental and sentence-level features, while only three felt they had an increased awareness of their word stress proficiency. For Group 2 the responses were more varied with seven Speakers (9, 10, 11, 12, 14, 15 and 16) stating increased self-awareness of individual sounds, four Speakers (9, 10, 11, 12 and 16) stating awareness of word stress, six speakers stating awareness of sentence-level features (Speakers 9, 10, 11, 12, 13 and 14) and 4 stating greater awareness of fluency (Speakers 9, 12, 13 and 16). Since both groups received similar intervention in terms of awareness-raising activities (Lesson 1 and Lesson 10), it is to be expected that the two groups would respond in a similar way and this was broadly speaking the case. All speakers, with the exception of Speaker 13 felt they had a greater self-awareness of their pronunciation of individual sounds and this was perhaps a result of rater comments, as discussed in 5.1.1, largely being related to segmental features. It was also the area most commented on in peer feedback. Only three Speakers in Group 1

and five in Group 2 felt they had greater self-awareness of word stress, possibly because this was only mentioned once in the rater feedback, when one of the raters commented that Speaker 14 said ‘BERlin instead of BerLIN’ (see Section 5.1.1). It was also only the focus of one lesson (Group 1 - Lesson 4). The majority of participants (eight in Group 1 and six in Group 2) felt they had greater awareness of sentence-level features. There was little rater feedback on this feature, with intonation and rhythm occasionally being mentioned, but it did feature in the lessons for Group 1 (Lesson 4, 5 and 8) and Group 2 (Lesson 3, with lessons 7, 8 and 9 featuring work on sentence-level practice). Finally, four speakers in each group felt they had greater awareness of their fluency. This was a feature more often commented on by raters who talked of ‘hesitancy’ or mentioned the speaker was ‘a bit slow’. It was also commented on in teacher feedback. It was therefore the areas of individual sound and sentence-level features in which participants felt they had gained self-awareness in both groups, with word stress and fluency mentioned less often.

It is important to mention that the feedback received by the participants was largely based on rater and NS teacher feedback, since the participants were, on the whole, reluctant or unable to comment on their fellow group members. This may have been due to the medium of instruction, which was via Zoom, and is discussed further in Section 5.5. below. Since the feedback given to the participants about their accents was, therefore, based on the rater feedback from the Pre-Intervention tokens and NS teacher feedback, it might be reasonable to assume that any change in participant phonological self-awareness was in the direction of NS judgements. The End of Course Questionnaire (Appendix E) asked participants what features they had an increased awareness of, but it did not ask them to specify whether these were areas which they needed to work on, so it was not possible to check if specific features corresponded to NS comments. In order to be able to answer this question more effectively, participants could have been asked to complete the Accent Rating Questionnaire relating to their own accent, so that it could be compared with rater feedback. However, since they had, by this point, received rater and teacher feedback on their accent, it would be unclear whether they were reciting what they had been told about their accent in the feedback or whether they really had achieved greater phonological self-awareness. This would appear to be something of a Catch 22 situation with assessing the findings of intervention to into raising learner awareness to their own accent. Nevertheless, it was the case that six participants in Group 1 and six in Group 2 (Table 49) stated that they found work on evaluating their current accent (End of Course Feedback Questionnaire, Appendix E, Question 6) was useful in helping them work on pronunciation.

5.4.3 Conclusion – Question 4

In answer to the research question, ‘*can intervention influence the learners’ ability to evaluate their own spontaneous speech production to correspond to that of native speaker raters?*’, qualitative responses from the participants indicated that all sixteen felt that the course had given them an increased awareness of their accent in English. The majority indicated that they had an increased self-awareness of individual sounds and sentence-level features, with approximately half of each group citing increased self-awareness of word stress and fluency. Since the feedback they received was from the raters and the NS teacher, it was assumed that the greater self-awareness was in the direction of NS raters, but this was not proven quantitatively. The End of Course Feedback Questionnaire also provided interesting insights into learner reactions to the course. Six of the participants no longer stated SBE as their pronunciation goal with three of the participants now saying they were happy with their accent. 13 of the 16 participants said they felt the course had helped them to achieve accent modification, the exceptions being Speaker 5 saying they were happy with their accent, Speaker 16 appearing to have made a mistake since they ticked ‘yes’ and ‘no’. This left only Speaker 1 who did not feel that their accent had changed. Group 1 particularly mentioned change at the segmental level, while Group 2 highlighted both segmental and sentence-level change. Finally, there appeared little consensus on which activities had been the most useful, suggesting that a variety of activities should be used in PI. While some might find a structural approach more helpful, others may respond to a more strategy-based approach.

5.5 GENERAL DISCUSSION OF RESULTS

The findings for Question 1 back up the findings of previous research in the field which suggests that many learners may struggle to evaluate their own accent and show a tendency to over-estimate their performance in L2 speech (Kennedy & Trofimovich, 2017; Saito et al., 2020; Trofimovich et al., 2016). There are important implications for the PI classroom, namely that phonological self-awareness should be the first activity in any PI intervention (Couper, 2015; McGregor & Reed, 2018), since the learner needs to understand, “the difference between what we actually say ... and what we think we say” (Couper, 2015, p.421). It is difficult to imagine how progress can be made if the learner is unaware of or unrealistic about how their accent is being perceived by others. Also, as discussed in Section 5.4.1.6, work on evaluating current accent was judged useful by 12 of the 16 participants, six in each group, and was therefore judged the most useful activity overall.

Similarly, the findings for Question 2 back up a growing body of research which suggests that, for pronunciation, instruction is more effective than mere exposure in late learners (Lee et al., 2015; Saito,

2012; Thomson & Derwing, 2015a). It supports the notion of fossilization (Selinker, 1972) discussed in Section 2.3 suggesting that, for pronunciation at least, after a certain point in L2 learning, the L2 accent becomes impervious to further change in the absence of intervention. The findings suggest that the interventions were effective in disrupting fossilization. Section 3.3.1 looked at how Long (2000) proposed three types of SLA instruction, namely: Focus on Meaning (FonM), Focus on Forms (FonFS) and Focus on Form (FonF). This finding suggests that an FonM approach, which involves communicative activities without overt attention to specific structures, may be insufficient to promote ‘noticing’ (Schmidt, 2012) with this learner group, as discussed in Section 3.1. The bottom-up approach followed by Group 1 can be seen as a largely Focus on Forms (FonFS) approach, with the focus on specific forms, described by Ellis (2016) as “a structure-based approach ... where linguistic forms are taught directly and explicitly” (p.405). The work on voiceless dental fricative /θ/ and voiced dental fricative /ð/ contrast in Lesson 2, for example, involved an exercise to hear the difference between the sounds, followed by a production exercise at the word level and then at the word level.

However, the top-down approach followed by Group 2 does not appear to neatly fit into any of the types of instruction proposed by Long (2000) in Section 3.3.1. While it bears similarities with the FonF approach where the focus on language forms is secondary to the focus on meaning, the top-down approach to PI asks for a focus on *performance*, where the focus on language forms is secondary. Many of the suggested strategies such as mirroring, echoing and PPR proposed in Section 3.4, focus on copying rather than meaning. The top-down approach also has some similarities with the FonM approach, in activities such as roleplay which elicits spontaneous speech, but again, the focus is on performance not meaning. It is perhaps the case, therefore, that the top-down approach requires a fourth type of SLA instruction to be added to the typology, namely Focus of Performance (FonP), an approach specific to the teaching of pronunciation since, as proposed in Section 3.1, pronunciation acquisition is largely a question of skill acquisition.

Perhaps the most interesting findings from this research come from question 3. While top-down strategies have been compared to bottom-up strategies in some areas of SLA, such as listening (Orii-Akita, 2014; Yeldam, 2016) and vocabulary learning (Lee & Yoon, 2019; Moskovsky et al., 2015), as yet there appears to be no research comparing a bottom-up versus a top-down approach in a general PI programme. As discussed in Section 5.3, the findings of this study suggest that both intervention programmes were effective in achieving statistically significant gains in nativeness, which seemed to be a result of changes in segmental/word-level features and sentence-level features. This is an important finding since it suggests that both approaches are effective and that, therefore, perhaps a mixed approach would be most productive in the SLA classroom, since different approaches might resonate with different students. For example, it has already been suggested (Section 5.3.3.1) how

Participant 15, who was in Group 2 following a top-down approach, might have benefitted more from a bottom-up approach given that NNS features of his accent were judged to be largely segmental. For a busy language teacher, it is not possible to analyse the specific needs and preferences of each individual student in each class being taught, so it makes sense that a scattergun approach introducing different activities and strategies would be effective in reaching a maximum number of learners.

Finally, Question 4, which found that participants felt that the intervention helped to give them an increased awareness of their accent in English, supports existing research suggesting that activities such as listening to recordings of their own speech (Couper, 2015), receiving expert feedback (Lappin-Fortin & Rye, 2014) and receiving peer feedback (Miller & Ng, 1994; Tsunemoto et al., 2022) is effective in raising phonological self-awareness.

However, there are some important caveats to these findings. Firstly, the research was carried out with a limited number of participants. While initially there were 32 participants, 16 in each group as discussed in Section 4.3, eight students dropped out and five did not attend a minimum of eight of the ten classes, so were dropped from the study. Unfortunately, four of the drop-outs were from Group 1 meaning that Group 1 had only eight participants. This resulted in only eight of the eleven participants from Group 2 being included in the findings in order to balance the groups. One problem with a small cohort can be seen with the rater scores for Speaker 15. It was suggested in Section 5.2.4.1 that the rater judgements for this speaker were something of an outlier compared to the general trend of Group 2 and this was further mentioned in Section 5.3. Table 50 shows the change in mean rater score from Pre- to Delayed Post-Intervention for Groups 1 and 2 and the figures for Group 2 when Speaker 15 is omitted.

Table 50. Change in Mean Rater Score from Pre to Delayed Post-Intervention per Measure for Group 1, Group 2 and Group 2 without Speaker 15.

Measure	Group 1 Change Pre To Delayed	Group 2 Change Pre To Delayed	Group 2 Change No Speaker 15 Pre To Delayed
Nativeness	1.53	0.86	1.18
Segment	1.13	0.82	1.04
Sentence	1.14	0.42	0.67
Fluency	1.10	-0.08	0.21
Vocab	0.69	0.04	0.30
Accuracy	0.31	0.15	0.35
Overall Change	5.90	2.21	3.75
Mean Change	0.98	0.37	0.63

As can be seen in Table 50, this reduces difference between Group 1 and Group 2 in terms of the mean rater score change for each measure and shows that, with a small cohort, one outlier can have a large effect on results. In order to obtain corroboration for the findings of this study and reduce individual participant effects, the study needs to be repeated with a larger cohort. As such, this study represents a preliminary finding in this area and the issue merits further study. Future research should consider using a larger cohort to carry out this and similar research.

It was also mentioned that attendance was set at a minimum of eight lessons. In fact, only one participant (in Group 1) attended only eight lessons, with three attending nine (2 from Group 1 and 1 from Group 2) with the remaining twelve participants attending all ten lessons. Even so, as mentioned by Participant 1 in the End of Course Questionnaire (Section 5.4.1.6), ten hours of intervention was relatively short and resulted in limited time for individual practice during lesson time. This was particularly so for Group 1, who had quite a lot of material to get through, listening to the feature of pronunciation which was the topic of the lesson and then carrying out controlled and then less controlled practice exercises. It is suggested that, in future research, a longer intervention period, perhaps allowing 90 minutes per lesson, and/or an increased number of lessons would address this problem.

In order to increase practice time, students were encouraged to do ‘homework’ and to work on the areas we had covered during each lesson, but it was unclear whether students were doing this. Indeed, the comment in the End of Course Questionnaire from Participant 11 mentioned in Section 5.4.1.6 – “*sadly I was very busy during the time the course took place, so I could no [sic] spend on it the time it deserved*” suggested that this participant was not completing homework tasks. After the first lesson, each lesson started with a review of the homework, with learners invited to show their progress or reflect on the task, but participants were generally reluctant to do this. This was possibly because they had not done the homework task, but it is also possible that participants had a degree of language anxiety (Section 2.6) and were reluctant to ‘perform’ in front of the class. It is also possible that, as all participants were either students or working adults, they had little time for extra practice. Future research should include a question about the extent of the participants’ out-of-class pronunciation practice in the End of Course Feedback in order to evaluate whether out-of-class practice was influencing outcomes.

It would also be useful to survey the participant at the time of the Delayed Post-Intervention regarding the issue of out-of-class practice. If, as has been argued, a large component of pronunciation modification is essentially skill acquisition (Section 3.1), accent is not something in which change would be expected overnight. Changing habits, which in late learners are possibly fossilized, (Section

2.3) takes time, practice, feedback and encouragement. Therefore, it would also have been interesting to ask participants at the time of the Delayed Post-Test, whether they had been working on their pronunciation in the three months following the intervention, particularly the ten participants who maintained SBE as their pronunciation goal discussed in Section 5.4.1.3, in order to see whether continued practice was influencing outcomes. Table 31 in Section 5.3.3.1 shows that there was very little mean change in rater scores for nativeness for Group 1 (0.01) or for Group 2 (-0.07) from the end of intervention (Post-Test) to three months after the end of the intervention (Delayed Post-Test), but a detailed look at individual speakers shows a more nuanced picture (Table 31). No participants made huge strides towards nativeness during this time period according to the raters, but there were some participants who showed change towards nativeness of over 0.5 of a point: Participant 6 (0.88), Participant 8 (0.55), Participant 14 (0.55) and Participant 16 (0.67). It would have been interesting to know whether this continued change in the direction of nativeness was due to continued effort on their part. Equally, there were those who showed a drop in their nativeness of over 0.5 of a point according to the raters: Participant 4 (-0.89), Participant 7 (-0.89), Participant 10 (-1.22) and Participant 15 (-1.22). Speakers 7 and 15 stated they were happy with their current accent, so it can be assumed they did not work further on accent modification, but it would be interesting to know if Speakers 4 and 10 also ceased to work on accent modification at this time. Future research should aim to question participants on whether they continued to work on their pronunciation between the Post- and Delayed Post-Intervention and if so, how much, in order to establish whether this was influencing outcomes in the Delayed Post-Test.

It is possible that student reluctance to ‘perform’ in front of the class was related to the fact that the classes were delivered on Zoom. Zoom lessons tend to be quite intense with all participants able to see each other at all times. While it was possible to divide students into Breakout Rooms to work on exercises in pairs or threes, it was time-consuming for the teacher to visit each group in turn in order to give them support and feedback. Working in smaller groups, in my experience, generally gives students more confidence. This aspect of group work would have been facilitated by face-to-face interactions, where the teacher could listen to groups in a less intrusive way and give rapid individual feedback. While initially the lessons were planned as face-to-face, the Covid pandemic of 2020-2021 meant that this decision had to be rethought. The pandemic was largely over by the Spring of 2022 when the intervention was carried out, but it was nevertheless decided that online lessons would be a more judicious option. Zoom lessons were also not ideal as the classes, particularly for Group 1, involved a lot of listening activities, either listening to recordings highlighting the features being discussed or listening to fellow participants in order to give them feedback on their performance of a given task. Clearly, listening through a computer audio system does not give the clarity that face-to-

face interaction gives, not to mention occasional problems with participants' screens freezing or lagging. It is possible, therefore, that, for the reasons suggested above, superior results may have been obtained had the classes been delivered face-to-face instead of online. While not always possible, it is suggested that future research should consider face-to-face classes for PI research.

There were methodological problems with comparing data regarding the self-assessment of accent on the Learner Consent Form and Pronunciation Questionnaire (Appendix D) with rater judgements in the Accent Rating Questionnaire (Appendix C). The raters were asked to listen to a speech token and rate the accent using a 9-point Likert scale. It was felt that the participants could not be asked to do this, since the act of listening to their speech token would already constitute an intervention and would therefore not be an accurate reflection of their Pre-Intervention judgements of their accent. Instead, they were asked to tick yes/no whether they had problems with the specific measures (Question 8). It is questionable whether these competing measures can be compared. As suggested in Section 5.1.3, a possible solution would have been to ask the participants to evaluate their accent using the Accent Rating Questionnaire, but without asking them to listen to the speech token. This would have enabled a more accurate comparison of rater and participant perceptions. Future research should consider using the same questionnaire for raters and participants, omitting the speech token for the participants. This procedure could then have been repeated in the End of Course Feedback in order to evaluate if participants had changed their self-assessment in the direction of the rater judgements. However, as discussed in Section 5.4.2, by this time the participants had received feedback from the teacher and raters on their accent, so it would be somewhat unclear as to whether participants were repeating what they had been told about their accent in the feedback or whether they really had achieved greater phonological self-awareness. It was therefore felt that qualitative data, asking participants whether they felt they had greater self-awareness, would be more revealing at this stage. A possible solution would be to ask participants to provide both quantitative data, in the form of the Accent Rating Questionnaire and qualitative data, in the form of the End of Course Feedback Questionnaire at the Post-Intervention time point, in order to achieve a more complete picture of participants' pronunciation self-awareness after the intervention. Future research should consider using both quantitative and qualitative data in the post intervention participant self-evaluation.

A further factor which needs to be taken into account with the findings, as discussed in Section 5.1.2, is the differences between the raters. While inter-rater reliability scores were judged to be good or very good using Cronbach's alpha (Section 5.2.1), it was noticeable that, in many cases, there was a marked difference of opinion between the raters. For example, in Section 5.1.1, Table 5 shows that Speaker 3 was judged as having an American, a Brazilian, an Iberian plus British and a British influence on their accent. As explained in Section 4.2.1, the raters were deliberately chosen to provide variety:

American/British; familiarity with Portuguese accent/no familiarity; experience in language teaching or linguistics/no experience and ages ranging from 25 to 71 at the start of the study. This means that the raters were not a homogeneous group and would appear to be responding differently to the different speakers. It would be informative to know what rater features were influencing these differences. For example, it is possible that American raters would have greater sensitivity to and therefore be more critical of, an American English influenced speaker and vice-versa with a British English influenced speaker. It is also possible that those with a familiarity with a Portuguese accent in English would be more tolerant of NNS features of the accent and rate more leniently, as proposed in Section 2.4.1.4. However, as this was not the goal of this research, rater feedback was given anonymously, so information relating to the influence of rater background on feedback was not obtained. In this research, it is possible that clearer findings would have been obtained had the raters been a more homogeneous group, for example, all experienced NSs of British English. However, as discussed in Section 5.2.1, there were valid reasons for not doing this and following the advice of Piske et al. (2001), that “a representative sample of raters should be recruited and not one particular type of rater” (p.195). Research into the influence of rater profile on rater judgements of both NNS and NS tokens would be an interesting and important avenue for future research.

It could also be the case that clearer findings would have been obtained had a more controlled token been elicited, for example, from a delayed repetition task or a read-aloud task. While the rating of spontaneous tokens is more likely to be influenced by lexicogrammatical errors as discussed in Section 4.2.2.3, the intention of this research was to help learners modify their accent in spontaneous speech, rather than in controlled and artificial tasks which are unrepresentative of real-life language use. In Section 3.3.4, it was stated that Thomson and Derwing (2015a) in their review of 75 L2 pronunciation studies, noted that only 20% of studies included spontaneous speaking tasks and that, within that group, some of them allowed time for preparation, which meant calling them ‘spontaneous’ was questionable. It is important to be aware that unless research can show that an intervention, whatever it may be, transfers to spontaneous speech production, it is of limited use to educators and can only claim to be of academic interest. This is an important consideration for future research into PI.

It may also be the case that the findings for Question 3 were clouded by a certain degree of overlap between the two interventions, but it is difficult to see how this could be avoided. Pennington (2019a) advocates that a top-down pronunciation approach, “would start at the level of broad social context, suprasegmentals, and related behaviours, then systematically progress ‘downwards’ to the segmental level” (p.381). The top-down approach, therefore, does not advocate the removal of a structural, suprasegmental and segmental approach, it proposes that these should be secondary to work helping the learner to build an awareness of themselves within their social environment. Pennington (ibid.)

describes how pronunciation modification is not just a matter of working on auditory perception and the motor ability to produce sounds, but also “a matter of openness and willingness to change based on valuing such change” (p.375). The Group 2 approach aimed to follow this procedure by first inviting learners to think about their current L2 pronunciation (Lesson 1) and then think about why they wanted to change and how they wanted to change (Lesson 3). This is in line with the proponents of student-centred learning (Section 3.4.1), such as Bernaus (2010) who comments, “if learners feel that they are their own masters in their learning they will be motivated to learn and of course to attain higher levels of achievement” (p.185). Group 1, while they were encouraged to think about their current L2 pronunciation (Lesson 1), were not invited to think about why they were not happy with it or why they wanted to modify their accent towards SBE. They just followed a structural syllabus which looked at possible problem areas for L1 Portuguese speakers of L2 English. Thus, while both groups looked, for example, at the tense/lax vowel distinction with focus on the front, high, lax vowel /ɪ/ (live) and the front, high, tense vowel /i:/ (leave), Group 1 (Lesson 2) approached it as a discrete structural comparison following a PPP approach (Section 3.3.3.1), while Group 2 (Lesson 4) looked at this as a strategy for practising and modifying a skill. Similarly, L1 interference was presented to Group 1 (Lesson 5) in terms of structures which they may need to work on using the proposed PPP framework. In Group 2 (Lesson 7), it was addressed after the lesson on aptitude and identity (Lesson 6), which suggested that modifying L2 pronunciation may involve letting go of the L1 self (Section 3.4.4) and embracing the ‘ideal L2 self’ (Dörnyei, 2009, Section 2.7.1).

Thus, while there was a limited amount of overlap between the two groups, the issues were addressed from different perspectives. It is also difficult to imagine how this could be avoided since, as mentioned above, top-down instruction does not preclude segmental and suprasegmental work, it merely states that these should be looked at within the bigger picture of social context. A possibility would be to limit the top-down instruction procedure to only look at strategies for working on fluent speech, such as shadowing, mirroring and the template and comparing this with a structural approach. This would, however, no longer be a comparison of top-down versus bottom-up approaches, but rather a comparison of structure-led versus strategy-led approach. This would be an interesting avenue for future research.

Another possible area of overlap between the two groups resulted from the difficulty of maintaining clear distinction between the approaches at all times. The intervention was not, and should not be, purely a case of the teacher lecturing the learners. The learners were invited, at every stage, to reflect on their learning and to give their input which, on occasion, led to the intervention going in an unintended direction. For example, at one point, a member of Group 1 asked whether it might be a good idea to try to copy the speech of a target speaker. While this was clearly a top-down, rather than

a bottom-up approach, the author found it impossible to discourage this idea (not to mention ethically questionable), which probably would have been helpful to the learners. It was accepted that this approach might be fruitful, but it was not encouraged or practised in the Group 1 lessons. It is very difficult to see how such an overlap caused by a learner contribution could be avoided. In the language classroom, it is important to create an atmosphere where the learners feel safe to ask questions and express themselves, particularly in a pronunciation class. Nevertheless, it has to be noted as a possible area of overlap between the groups and this may have contributed to the lack of a significant finding between them.

To summarise, while this research supports previous findings into a possible learner difficulty with phonological self-assessment, the success of PI over exposure in late learners and the effectiveness of activities to raise learner awareness to their own accent, the most interesting finding possibly came from Question 3. It would appear that, in the field of PI, comparing top-down versus bottom-up approaches to accent modification has not yet been addressed in the research. This study suggested that both were effective.

However, there were some caveats to the findings. Firstly, a larger cohort of participants is needed in order to reduce outlier effects and ascertain whether the findings can be replicated with a larger group of participants. Secondly, the intervention period of ten hours was relatively short and it is likely that clearer findings could be obtained with a longer course of intervention. It would also be useful to question participants both in the End of Course Feedback Questionnaire and on collection of the Delayed Post-Intervention token about their out-of-class practice during the intervention and between the Post-Intervention and Delayed Post-Intervention time points respectively, in order to establish whether out-of-class practice was influencing outcomes. While this may not always be possible, it is proposed that, in the case of PI, face-to-face classes are preferable to online lessons due to the intense nature of online lessons and to possible audio-visual issues with online platforms. Using raters from a variety of different backgrounds provides a more representative sample of the listeners with whom the participants are likely to come into contact in their real-world interactions, but it would also be interesting to carry out research into the influence of rater profile on judgements of NS and NNS tokens. Using spontaneous speech tokens, while problematic as they are more likely to be subject to the influence of lexicogrammatical errors, is nevertheless fundamental in informing classroom practice and helping learners in their language use outside the classroom. Lastly, while this may not always be possible, it is important to avoid overlap between intervention groups as much as possible. While this is problematic with a comparison of top-down versus bottom-up instruction, since top-down does not preclude work on segmental and suprasegmental features, a comparison of structure-led versus strategy-led approach might provide a clearer distinction and be a fruitful avenue for future research.

5.6 SUMMARY OF FINDINGS

In Question 1, it was predicted that, in line with the predictions of the Dunning-Kruger effect (Section 2.4.1.4), many learners would judge their accents more favourably than would NS raters and this was supported by the results of this study, although there appears to be a certain amount of individual variation in participant phonological self-awareness suggesting it is an aspect of language aptitude. If this is indeed the case, raising learner phonological self-awareness would appear to be an important starting point for any PI programme.

In terms of Question 2, intervention versus exposure, both interventions were found to be superior to the Control in modifying participant accents. A significant interaction was found between time and group for nativeness, which means which group a participant was in influenced their nativeness results. This appeared to be due to a significant increase in sentence-level scores for the intervention groups. For segmental/word-level features, there was greater increase in the intervention groups than in the Control, although this was not statistically significant. No statistically significant difference between the groups was found for fluency, vocabulary and accuracy, which is to be expected as these areas were not the focus of the intervention.

Question 3 involved a comparison of bottom-up versus top-down intervention and both groups saw changes towards nativeness according to rater judgements. The bottom-up approach was superior to the top-down in terms of fluency, but this was possibly due to Pre-Intervention group age differences. Participants varied in which activities they found most useful in helping them to achieve accent modification. This suggests that a scattergun approach to teaching pronunciation might be the most effective, where both top-down *and* bottom-up approaches are used in the hope that, while one activity may not lead to change in a given learner, another one might.

Finally, Question 4 asked if intervention could raise participant phonological self-awareness. Qualitative responses from the participants indicated that all sixteen felt that the course had given them an increased awareness of their accent in English. The majority indicated that they had an increased self-awareness of individual sounds and sentence-level features, with approximately half of each group citing increased self-awareness of word stress and fluency. Since the feedback they received was from the raters and the NS teacher, it was assumed that the greater self-awareness was in the direction of NS raters.

However, caveats to the findings relating to cohort size, length of intervention, amount of out-of-class practice, mode of instruction (online not face-to-face), methodological problems of participant/rater comparison, nature of the token (spontaneous), profile of raters and degree of overlap between the two

intervention approaches mean the further research is needed in order to provide support for conclusions drawn in this study.

CONCLUSION

The objective of this study was to inform PI methodology in the L2 classroom, with a particular focus on the role of phonological self-awareness and the comparison of a bottom-up versus a top-down approach to instruction.

It was argued that NS and NNS are not discrete terms and that there is a ‘shadowy borderline’ (Escudero & Sharwood Smith, 2001, p.277) where the two overlap, meaning that the two terms can best be defined in terms of listener judgements of accent (Davies, 2003; Derwing & Munro, 2009). Escudero and Sharwood Smith (ibid.) use Prototype Theory (Rosch, 1975, 1978) to explain how certain aspects of language performance in general, and pronunciation in particular, might be judged more fundamental or *prototypical* to the concept of the NS accent, while other aspects might be seen as more peripheral and therefore less clear-cut in terms of nativeness judgements. It is the standard accents which probably lay claim to being most prototypical (Davies, 2003; Verbeke & Simon, 2023), although it has to be borne in mind that, even with standard accents, judgements are influenced by the listener’s familiarity with a given accent, by context and expectations (Verbeke & Simon, ibid.).

Over the last 30 years, attitudes towards the standard accent as the *target accent* of PI, the accent used as the pronunciation model in the classroom, have changed and important voices in the field of PI have argued for a move away from *nativeness* (sounding like an NS) as the goal, to *intelligibility* (getting the message across) and *comprehensibility* (the listener’s perception of difficulty in understanding the message) as the goal (Levis, 2020; Derwing & Munro, 2005). However, this change has resulted in a central problem for PI. If a speaker is deemed *unintelligible*, there must be some way of ‘correcting’ that and of offering an intelligible alternative. Lewis and Deterding (2018) sum up the centrality of this problem when they state, “if there is no clear model to refer to, how do we give guidance to students about how to improve their pronunciation, and how do we determine what needs to be fixed in order to enhance intelligibility?” (p.161). Since English is increasingly being used as the international language of communication in many domains of life, it is important that speakers have some degree of shared understanding of English in general and pronunciation in particular in order to be mutually intelligible. That this is an extremely serious issue, is highlighted by work such as Tiewtrakul and Fletcher (2009) who, in a review of communications between air-traffic control and pilots conclude that, “communication errors, defined by incidents of pilots not understanding, occur significantly more often when speakers are both non-native English” (p.229), and research by Lindhout et al. (2009), which estimated that language problems account for 5% to 10% of all accidents in the chemical industry. The point is, even if the *goal* of pronunciation instruction is intelligibility/comprehensibility, there still has to be a *target* accent, which provides the model for correction if intelligibility/comprehensibility is

compromised. As previously argued, it is the standard accents which probably lay claim to being most prototypical (Davies, 2003; Verbeke & Simon, 2023) and therefore, the most widely understood. This makes them strong candidates to be the target accent, even in intelligibility/comprehensibility-focused PI. However, ultimately, it is the learner who needs to make this decision. Bernaus (2010) comments, “if learners feel that they are their own masters in their learning they will be motivated to learn and of course to attain higher levels of achievement” (p.185). Accent, perhaps more than other areas of SLA, cannot be imposed upon a learner. Accent is a reflection of personal and group identity and therefore, as Pennington (2019a) points out, “...changing pronunciation means altering psychological, social, and cultural aspects of the self” (p.375). Accent modification can only be achieved if the learner *wants* to alter these aspects of the self and research suggest that many learners *do* want to achieve this and choose nativeness rather than intelligibility as their goal (Derwing, 2003; Jenkins, 2007; Wach, 2011).

In view of the above discussion, the primary objective of this study was to look at ways to help learners work on accent modification in the direction of nativeness. As a first step, this involved an overview of current research findings into the factors influencing ultimate attainment in pronunciation in late learners. The CPH as it related to language learning proposes an age-related *window of opportunity* for language learning prior to puberty, after which time decreasing brain plasticity makes it increasingly difficult to acquire a language and maybe impossible to acquire native-like language competence (Lenneberg, 1967). However, neither the CPH nor the subsequent *Sensitive Period* (Long, 1990), and *Linear Increase* (Oyama, 1976) hypotheses, satisfactorily explain anomalies in phonological attainment, such as early learners who are judged to have a ‘foreign accent’ (Flege et al., 1997; Uzal et al., 2015) and so-called exceptional learners (Birdsong, 2007; Bongaerts et al., 1995; Dollman, 2020; Ioup et al. 1994; Moyer, 1999), who are judged to have native-like pronunciation despite a late start. Subsequent research has presented a more nuanced picture and involves factors beyond the age constraint, such as native language interference (Best, 1994; Escudero, 2009; Flege & Bohn, 2021;), quantity and quality of exposure to and use of the language (Piske et al., 2001; Uzal et al., 2015), language aptitude (Abrahamsson & Hyltensam, 2008; Sleva & Miyake, 2006), gender (Moyer, 2016; Van der Slik et al., 2015), motivation (Moyer, 2014), degree of phonological self-awareness (Trofimovich et al. 2016; Kennedy & Trofimovich, 2017), personality (Hu & Reiterer, 2009; Lewandowski & Jilka, 2019) and identity (Gatbonton et al., 2005).

Importantly for this research, while some of these factors are immutable (age of onset, gender, personality) others have been shown to be modifiable (language exposure, negative L1 transfer, phonological self-awareness, motivation, identity and, possibly, cognitive control) and nowhere have these factors been shown to be insurmountable given the appropriate intervention. This study particularly focused on the area of phonological self-awareness.

It has been suggested that learners might struggle to evaluate their own accent and show a tendency to over-estimate their performance in L2 speech (Kennedy & Trofimovich, 2017; Saito, Trofimovich, Abe & In'nami, 2020; Trofimovich et al., 2016), while good self-monitoring skills appear to be a feature of exceptional learners (Moyer, 2018; Muñoz & Singleton, 2011). Researchers have argued that phonological self-awareness needs to be one of the first steps of any intervention to modify pronunciation (Couper, 2015; McGregor & Reed, 2018; Sardegna & McGregor, 2013) and that research has shown success in helping learners to gain self-awareness (Lappin-Fortin & Rye, 2014; MacDonald, 2018; Tsunemoto et al., 2022). This study therefore hoped to build on this research by comparing learner descriptions of their own accent with NS ratings of their Pre-Intervention speech token. The study also looked at whether self-awareness could be raised by certain activities: having learners record and listen to a passage of their own speech (Couper, *ibid.*; McDonald, *ibid.*) and receiving expert feedback (Lappin-Fortin & Rye, *ibid.*) and peer feedback (Tsunemoto et al. *ibid.*). It was expected that receiving pronunciation feedback from the instructor, the raters and other learners in the group, along with encouraging participants to record and listen to their own output, would give them greater awareness of their current accent.

The findings indicated that, in this cohort, 25% of the participants produced judgments of their accent which were close to that of the NS raters, but the remaining 75% were either unable to describe their accent or described it differently from the NS raters. The fact that the participants chose to take part in a pronunciation course and that several of them worked in the area of language acquisition suggested that this cohort would, in fact, be *more* self-aware than the average learner, which would imply that the actual percentage of learners who are able to match NS judgements of their accent may be even smaller than the 25% identified here. Furthermore, qualitative responses from the End of Course Feedback Questionnaire (Appendix E) indicated that all sixteen participants felt that the course had given them an increased awareness of their accent in English. The majority indicated that they had an increased self-awareness of individual sounds and sentence-level features, with approximately half of each group citing increased self-awareness of word stress and fluency. Since the feedback they received was largely from the raters and the NS teacher, it was assumed that the greater self-awareness was in the direction of NS raters.

This study also looked at research findings into the influence of PI on accent modification. Thomson and Derwing (2015a), in a review of 75 L2 PI studies, reported that 82% of the research they surveyed reported successful outcomes, but noted that only 20% of studies included spontaneous speaking tasks and that, within that group, some of them allowed time for preparation, which meant calling them 'spontaneous' was questionable. More work needs to be done to establish whether production gains at the segmental and suprasegmental level described above can transfer to spontaneous speech.

Meanwhile, research is beginning to suggest that intervention might be more effective if students are encouraged to go beyond a focus on specific features of pronunciation, be they segmental or suprasegmental, and to think more holistically about PI in terms of helping the learners “to develop awareness of themselves as speakers ...” (MacDonald, 2018, p.249). LaScotte et al. (2021) describe how the research and teaching of L2 PI largely falls into two groups: the ‘bottom-up’ approach, which is essentially structural, focusing on de-contextualised phonological forms, and the ‘top-down’ approach in which communication and the social context are paramount (Pennington, 2019). In the top-down paradigm, work on phonological structures is secondary to strategies such as shadowing (Foote & McDonough, 2017; Martinsen et al, 2017), echoing (Chung, 2017), mirroring (Meyers, 2016), channeling (La Scotte & Tarone, 2019; Moreno, 2016) and roleplay (Pennington & Waxler, 2017), which aim to work on pronunciation holistically. Research has been done to compare the relative effectiveness of a bottom-up versus a top-down approach in the area of SLA listening (Orie-Akita, 2014; Yeldam, 2016) and vocabulary learning (Lee & Yoon, 2019; Moskovsky et al., 2015), but as yet there appears to be no research comparing a bottom-up versus a top-down approach in a general PI programme. The current study hoped to begin to address this lacuna by comparing the two approaches. It was predicted that the top-down approach, with its more holistic and communicative-led focus, would produce superior results to a purely structural syllabus in line with the recommendations of prominent researchers (LaScotte et al, *ibid.*; Moyer, *ibid.*; Pennington, *ibid.*) and that both would be superior to exposure alone, as predicted by the Noticing Hypothesis.

In terms of intervention versus mere exposure, both intervention groups were found to be superior to the Control in modifying participants’ accents. A significant interaction was found between time and group for nativeness, which means the group a participant was in influenced their nativeness results. This appeared to be due to a significant increase in sentence-level scores for the intervention groups. For segmental/word-level features, there was greater increase in the intervention groups than in the Control, although this was not statistically significant. No statistically significant difference between the groups was found for fluency, vocabulary or accuracy, which is to be expected as these areas were not the focus of the intervention.

Regarding the comparison between the bottom-up versus top-down intervention, both groups saw changes towards nativeness according to rater judgements. There were significant changes in mean rater score for segmental/word-level features and sentence-level features in the direction of nativeness, but no significant difference between the two groups. The bottom-up approach was superior to the top-down in terms of fluency, but this was possibly due to group age differences, as the mean age of Group 1 (31.1 years) was over 6 years greater than the mean age of Group 2 (24.8 years). Vocabulary and accuracy measures appeared to be unchanged by the intervention and this was to be expected as the

interventions did not target these areas. Participants varied in which activities they found most useful in helping them to achieve accent modification, which suggests that a ‘scattergun’ approach to teaching pronunciation might be the most effective, where both top-down *and* bottom-up approaches are used in the hope that, while one activity may not lead to change in a given learner, another one might. This would be an interesting area for future study.

While this study suggested that both interventions were effective in achieving change in NS rater scores in the direction of nativeness, there were some caveats to the findings. Firstly, a larger cohort of participants should have been used in order to reduce outlier effects and ascertain whether the findings could be generalised with a larger group of participants. Secondly, the intervention period of ten hours was relatively short and it is likely that clearer findings could have been obtained with a longer course of intervention, possibly increasing lessons to 1.5 hours or extending the duration of the course beyond the ten lessons. It would also have been useful to question participants both in the End of Course Feedback Questionnaire and on collection of the Delayed Post-Intervention token about their out-of-class practice during the intervention and between the Post-Intervention and Delayed Post-Intervention time points respectively, in order to establish whether out-of-class practice had influenced outcomes. This is important since Skill Acquisition Theory proposes that considerable practice is needed to achieve automatization of procedural knowledge (DeKeyser & Criado; 2013; DeKeyser, 2015). It was also felt that, in the case of PI, face-to-face classes would have been preferable to online lessons due to the intense nature of online lessons and to possible audio-visual issues with online platforms. There were methodological problems with comparison of participant and rater perceptions of accent and it is important, going forward, that the same rating questionnaire is used by both groups. Using raters from a variety of different backgrounds and using spontaneous speech tokens was felt to have diluted the findings to some extent, in that it probably led to decreased inter-rater reliability. However, using a representative sample of the listeners with whom the participants are likely to come into contact in their real-world interactions, and testing whether intervention activities influence spontaneous speech production are fundamental in informing classroom practice and helping learners in their language use outside the classroom. Lastly, there was some degree of overlap between intervention groups, since top-down instruction does not preclude work on segmental and suprasegmental features. However, a comparison of a structure-led versus a strategy-led approach might provide a clearer distinction and be a fruitful avenue for future research.

In conclusion, this study attempted to distil what research tells us about L2 pronunciation acquisition and transfer it to a classroom context. Methodological limitations may have diluted the results, but the findings were promising. Activities to raise phonological self-awareness seemed to be successful and both the intervention groups showed accent modification, which was not present in the Control Group

and this change maintained three months after the intervention. Perhaps the most useful finding was that participants differed in the tasks they found useful in working towards accent modification. This should not be a surprise given that the study involved a group of learners, all of whom are unique and respond differently to classroom tasks, but the implication is that the teacher should be using a variety of activities and strategies to work on pronunciation modification. Some learners may respond to segmental training, while others prefer suprasegmental work or top-down strategies such as roleplays and mimicry – variety would seem to be the key.

A broader approach to pronunciation instruction is not a new idea. Experts in this field have been arguing for the inclusion of a more top-down approach to pronunciation instruction for some time (LaScotte et al, *ibid.*; Moyer, *ibid.*; Pennington, *ibid.*), and for expanding the more traditional bottom-up approach to include work on the social and psychological aspects of accent. This research provides evidence in favour of this idea and adds one small piece to the puzzle of PI by suggesting that it is now time for top-down strategies to move from the theoretical to the practical, from the research paper to the classroom: segmental, suprasegmental and beyond.

REFERENCES

- Abrahamsson, N., & Hyltenstam, K. (2008). The robustness of aptitude effects in near-native second language acquisition. *Studies in second language acquisition*, 30(4), 481-509.
- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59, 249 – 306.
- Adrián, M. M. (2014). The efficacy of a reading aloud task in the teaching of pronunciation. *Journal of English Studies*, 12, 95-112.
- Alves, U. K., & Magro, V. (2011). Raising awareness of L2 phonology: Explicit instruction and the acquisition of aspirated/p/by Brazilian Portuguese speakers. *Letras de hoje. Porto Alegre*. 46, 3, 71-80.
- Anderson-Hsieh, J., & Koehler, K. (1988). The effect of foreign accent and speaking rate on native speaker comprehension. *Language Learning*, 38(4), 561-613.
- Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The relationship between native speaker judgments of nonnative pronunciation and deviance in segmental, prosody, and syllable structure. *Language learning*, 42(4), 529-555.
- Appel, R., Trofimovich P., Saito, K., Isaacs, T., & Webb, S. (2019). Lexical aspects of comprehensibility and nativeness from the perspective of native-speaking English raters. *International Journal of Applied Linguistics*, 170(1), 24–53.
- Archibald, J. (2018). Advanced-level L2 phonology. *The handbook of advanced proficiency in second language acquisition*, 239-263.
- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation*, vol VIII. G. Bower.
- Baker, A. (2014). Exploring teachers' knowledge of second language pronunciation techniques: Teacher cognitions, observed classroom practices, and student perceptions. *Tesol Quarterly*, 48(1), 136-163.
- Baran-Lucarz, M. (2016). Conceptualizing and measuring the construct of pronunciation anxiety: Results of a pilot study. *Classroom-oriented research: Reconciling theory and practice*, 39-56.
- Baran-Lucarz, M. (2023). Personality as a Correlate of Accentedness: The Case of Formal Setting Without Pronunciation-Focused Instruction. *Research in Language (RiL)*, 21(4), 377-396.
- Barrios, S., Jiang, N., & Idsardi, W. J. (2016). Similarity in L2 Phonology: Evidence from L1 Spanish Late-Learners' Perception and Lexical Representation of English Vowel Contrasts. *Second Language Research*, 32(3), 367–395.
- Bassetti, B. (2008) Orthographic input and second language phonology. In Piske, T. and YoungScholten, M. (eds.), *Input Matters in SLA*. Clevedon, UK: Multilingual Matters, 191-206.
- Bent, T., & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *The Journal of the Acoustical Society of America*, 114(3), 1600-1610.
- Bernaus, M. (2010). Language classroom motivation. In R. C. Gardner (ed.), *Motivation and second language acquisition: The socio-educational model*, 181–199.
- Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words*, 167(224), 233-277.

- Best, C., & McRoberts, G. (1989). Phonological influence on infants' perception of two nonnative speech contrasts. Paper presented at the *Society for Research in Child Development*, Kansas City, MO.
- Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. *Second language acquisition and the critical period hypothesis*, 161-181.
- Biedroń, A., & Szczepaniak, A. (2009). The cognitive profile of a talented foreign language learner. A case study. *Psychology of Language and Communication*, 13(1), 53-71.
- Biedroń, A., & Szczepaniak, A. (2012). Working memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal*, 96(2), 290-306.
- Birdsong, D. (1999). *Second language acquisition research. Second language acquisition and the Critical Period Hypothesis*. Lawrence Erlbaum Associates Publishers.
- Birdsong, D. (2006). Age and Second Language Acquisition and Processing: A Selective Overview. *Language Learning*, 56, 9-49.
- Birdsong, D. (2007) Nativelike pronunciation among late learners of French as a second language. In O.-S. Bohn & M. J. Munro (eds.), *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*. John Benjamins, 99-116.
- Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition*, 41-68.
- Boberg, C. (2015). North American English. In Reed, M. & Levis, J. (Eds.), *The Handbook of English Pronunciation*, 229 – 249.
- Bongaerts, T., Planken, B., & Schils, E. (1995). Can late starters attain a native accent in a foreign language? A test of the critical period hypothesis. In D. Singleton & Z Lengyel (Eds), *The age factor in second language acquisition*, Clevedon: UK, Multilingual Matters, 30-50.
- Bongaerts, T., 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.
- Borrigan, M., Martin, C. D., de Bruin, A., & Duñabeitia, J. A. (2018). Exploring Different Types of Inhibition During Bilingual Language Production. *Frontiers in psychology*, 9, 225
- Bowen, J. D. (1972). Contextualizing pronunciation practice in the ESOL classroom. *TESOL Quarterly*, 83-94.
- 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.
- Brennan, E. M., & Brennan, J. S. (1981a). Accent Scaling and Language Attitudes: Reactions to Mexican American English Speech. *Language and Speech*, 24(3), 207-221.
- Brinton, D. M. (2017). Innovations in pronunciation teaching. In *The Routledge handbook of contemporary English pronunciation*. 448-461).
- Buckingham, L. (2014). Attitudes to English Teachers' accents in the Gulf. *International Journal of Applied Linguistics*, 24, 50-73.
- Butler, J., & Frota, S. (2018). Emerging word segmentation abilities in European Portuguese-learning infants: new evidence for the rhythmic unit and the edge factor. *Journal of child language*, 45(6), 1294-1308.

- Cai, Y. W., & Lee, H. Y. H. (2015). A contrastive analysis of Mandarin Chinese and Thai: Suggestions for second language pronunciation. *Theory and Practice in Language Studies*, 5(4), 719-728.
- Campos, A. D., Mendes Oliveira, H., & Soares, A. P. (2018). The role of syllables in intermediate-depth stress-timed languages: Masked priming evidence in European Portuguese. *Reading and Writing*, 31, 1209-1229.
- Cao, R., & Jin, S. (2017). Phonological differences between received pronunciation and standard Scottish English. In *2017 International Seminar on Social Science and Humanities Research (SSHRR 2017)*, 121-130, Atlantis Press.
- Carlet, A., & Souza, H. K. D. (2018). Improving L2 pronunciation inside and outside the classroom: Perception, production and autonomous learning of L2 vowels. *Ilha do Desterro*, 71, 99-123.
- Carroll, J. B., & Sapon, S. M. (1959). Modern language aptitude test.
- Carter, T. J., & Dunning, D. (2008). Faulty Self-Assessment: Why Evaluating One's Own Competence Is an Intrinsically Difficult Task. *Social and Personality Psychology Compass*, 2, 346-360.
- Castelo, A. (2018). Ensino da componente fonético-fonológica: uma síntese e um exemplo de português para estrangeiros. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto*, 12, 41-72.
- Celce-Murcia, M. (1983). "Teaching pronunciation communicatively." *MEXTESOL Journal*, 7 (1), 10-23.
- Celce-Murcia, M. (1987). "Teaching pronunciation as communication." In J. Morley, ed. *Current Perspectives on Pronunciation: Practices Anchored in Theory*. Washington, D. C. TESOL. 5-12.
- Celce-Murcia, M. (1994). Teaching Pronunciation for Communication: The second time around. *MEXTESOL Journal*, 18(2), 25-34.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching Pronunciation: A Course Book and Reference Guide*. Cambridge University Press.
- Chen, X., He, J., Swanson, E., Cai, Z., & Fan, X. (2021). Big five personality traits and second language learning: A meta-analysis of 40 years' research. *Educational Psychology Review*, 1-37.
- Christiner, M., & Reiterer, S. M. (2013). Song and speech: examining the link between singing talent and speech imitation ability. *Frontiers in psychology*, 4, 874.
- Chung, K. S. (2017). Teaching pronunciation to adult learners of English. In H. Widodo, A. Wood, D. Gupta, & W. Cheng (Eds), *Asian English language classrooms: Where theory and practice meet*, 131-149.
- Clare, A., & Wilson, J. J. (2022). *Speakout: pre-intermediate students' book with activebook*, Pearson Education.
- Clément, R., Dörnyei, Z., & Noels, K. A. (1994). Motivation, self-confidence, and group cohesion in the foreign language classroom. *Language learning*, 44(3), 417-448.
- Collins, B., & Mees, I. M. (2009). *Practical phonetics and phonology: A resource book for students*. Routledge.
- Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. (2001). Council of Europe, Council for Cultural Co-operation, Education Committee, Modern Languages Division. Cambridge University Press.

- Common European Framework of Reference for Languages Learning Teaching Assessment. Companion Volume with new descriptors. (2020). Council for Cultural Co-operation, Education Committee, Modern Languages Division. Cambridge University Press.
- Corder, S. P. (1967). The significance of learner's errors. *IRAL: International Review of Applied Linguistics in Language Teaching*.
- Costa Jr, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of personality disorders*, 6(4), 343-359.
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of memory and Language*, 50(4), 491-511.
- Coumel, M., Groß, C., Sommer-Lolei, S., & Christiner, M. (2023). The Contribution of Music Abilities and Phonetic Aptitude to L2 Accent Faking Ability. *Languages*, 8(1), 68.
- Couper, G. (2003). The value of an explicit pronunciation syllabus in ESOL teaching. *Prospect* 18 (3), 53-70.
- Couper, G. (2006). The short and long-term effects of pronunciation instruction. *Prospect*, 21(1), 46-66.
- Couper, G. (2011). What makes pronunciation teaching work? Testing for the effect of two variables: socially constructed metalanguage and critical listening. *Language Awareness*, 20(3), 159-182.
- Couper, G. (2015). Applying theories of language and learning to teaching pronunciation. *The handbook of English pronunciation*, 413-432.
- Coyle, D. (2007). Content and language integrated learning: Towards a connected research agenda for CLIL pedagogies. *International journal of bilingual education and bilingualism*, 10(5), 543-562.
- Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2018) Linguistic Dimensions of Second Language Accentedness and Comprehensibility Vary across Speaking Tasks. *Studies in Second Language Acquisition*, 40(2), 443-457.
- Cruz-Ferreira, M. (1995). European Portuguese. *Journal of the International Phonetic Association*, 25(2), 90-94.
- Cruz-Ferreira, M. (1998). Intonation in European Portuguese. *Intonation systems. A survey of twenty languages*, 167-178.
- Cruz-Ferreira, M. (2002). Portuguese and English intonation in contrast. *Languages in Contrast*, 4(2), 213-232.
- Crystal, D. (1985). *A Dictionary of Linguistics and Phonetics*. Basil Blackwell.
- Crystal, D. (2003). *English as a Global Language*. Cambridge University Press.
- Crystal, D. (2006). *The fight for English: How language pundits ate, shot, and left*. Oxford University Press.
- Cunningham, S., Moor P. & Bygrave J. (2021). Cutting Edge Intermediate Student's Book & eBook with Online Practice, Digital Resources. Pearson.
- Dalton, C., & Seidlhofer, B. (1994). *Pronunciation*. Oxford University Press.
- Darcy, I. (2018). Powerful and Effective Pronunciation Instruction: How Can We Achieve It?. *CATESOL Journal*, 30(1), 13-45.
- Darcy, I., Mora, J. C., & Daidone, D. (2014). Attention control and inhibition influence phonological development in a second language.

- Darcy, I., Park, H., & Yang, C. L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, 40, 63-72.
- Davies, A. (2003). *The Native Speaker: Myth and Reality*. Clevedon: Multilingual Matters.
- Davies, A., & Elder, C. (2006). *The Handbook of Applied Linguistics*. Malden, MA: Blackwell Pub.
- Davis, R. F. (2019). *The effects of pitch and speaking rate on foreign accented speech perception*. BA Thesis, Louisville University.
- De Graaff, R., & Housen, A. (2009). Investigating the effects and effectiveness of L2 instruction. *The handbook of language teaching*, 726-755.
- DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. *Focus on form in classroom second language acquisition*, 28, 42-63.
- DeKeyser, R. (2015). Skill acquisition theory. In *Theories in second language acquisition* (Eds. VanPatten, B., & Williams, J.), 83-104.
- DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in second language acquisition*, 22(4), 499-533.
- Dekeyser, R. and Criado, R. (2013). Automatization, Skill Acquisition, and Practice in Second Language Acquisition. In *The Encyclopedia of Applied Linguistics*, C.A. Chapelle (Ed.).
- Derwing, T. (2003). What Do ESL Students say about Their Accents? *Canadian Modern Language Review-revue Canadienne Des Langues Vivantes*, 59, 547-567.
- Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in second language acquisition*, 19(1), 1-16.
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393-410.
- Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language learning*, 54(4), 655-679.
- Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39, 379-397.
- Derwing, T. M., & Munro, M.J. (2009). Putting Accent in Its Place: Rethinking Obstacles to Communication. *Language Teaching*, 42, 476-490.
- Derwing, T. M., & Munro, M.J. (2015). *Pronunciation Fundamentals: Evidence-based Perspectives for L2 Teaching and Research*. John Benjamins.
- Dewaele, J. M. (2005). Investigating the psychological and emotional dimensions in instructed language learning: Obstacles and possibilities. *The modern language Journal*, 89(3), 367-380.
- Díaz-Campos, M. (2004). Context of learning in the acquisition of Spanish second language phonology. *Studies in second language acquisition*, 26(2), 249-273.
- Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction*. Boston: Pearson/Allyn & Bacon.
- Ding, S., Liberatore, C., Sonsaat, S., Lučić, I., Silpachai, A., Zhao, G., Chukharev-Hudilainen, E., Levis, J. and Gutierrez-Osuna, R. (2019). Golden speaker builder—An interactive tool for pronunciation training. *Speech Communication*, 115, 51-66.

- Dollmann, J., Kogan, I., & Weißmann, M. (2020). Speaking accent-free in L2 beyond the critical period: The compensatory role of individual abilities and opportunity structures. *Applied Linguistics*, 41(5), 787-809.
- Dolman, M., & Spring, R. (2014). To what extent does musical aptitude influence foreign language pronunciation skills? A multi-factorial analysis of Japanese learners of English. *World Journal of English Language*, 4(4), 1.
- Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Dörnyei, Z. (2009). The L2 Motivational Self System. In Z. Dörnyei, & E. Ushioda (Eds.) *Motivation, Language Identity and the L2 Self*, Clevedon: Multilingual Matters.
- Doughty, C. (2019). Cognitive language aptitude. *Language learning*, 69, 101-126.
- Doughty, C., & Williams, J. (Eds.). (1998). *Focus on forms in classroom second language acquisition*. Cambridge, England: Cambridge University Press.
- Doughty, C., Campbell, S. G., Bunting, M. F., Bowles, A., & Haarmann, H. (2007). The development of the high-level language aptitude battery. *Center for Advanced Study of Language Technical Report (TTO 2105 M. 4)*. College Park, MD: University of Maryland.
- Dufour, S., Nguyen, N., & Frauenfelder, U. H. (2010). Does training on a phonemic contrast absent in the listener's dialect influence word recognition?. *The Journal of the Acoustical Society of America*, 128(1), 43-48.
- Edwards, I. G. H. (2008). Social factors and variation in production in L2 phonology. *Phonology and second language acquisition*, 36, 251.
- Elliott, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *Hispania*, 95-108.
- Ellis, R. (2006). Researching the effects of form-focussed instruction on L2 acquisition. *AILA Review* 19, 18-41.
- Ellis, R. (2009). Implicit and explicit learning, knowledge and instruction. In R. Ellis et al. (Eds.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching*. *Multilingual Matters*, 3-23.
- Ellis, R. (2016). Focus on form: A critical review. *Language teaching research*, 20(3), 405-428.
- Ellis, R., Basturkmen, H., & Loewen, S. (2002). Doing focus-on-form. *System*, 30(4), 419-432.
- English Style Guide*. (2016). Brussels: Commission of the European Communities.
- Escudero, P. (2005). *Linguistic perception and second language acquisition*. LOT Landelijke Onderzoekschool Taalwetenschap.
- Escudero, P. (2009). The linguistic perception of SIMILAR L2 sounds. In P. Boersma & S. Hamann (Ed.), *Phonology in Perception*, 151-190.
- Escudero, P., & Smith, M. S. (2001). Reinventing the native speaker: or 'What you never wanted to know about the native speaker so never dared to ask.'. *EuroSLA Yearbook*, 1(1), 275-286.
- Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics*, 36(2), 345-360.
- Eubank, L., & Gregg, K. R. (1999). Critical periods and (second) language acquisition: Divide et impera. *Second language acquisition and the critical period hypothesis*, 8(3), 65-99.

- Fergusson, G. (2009). Issues in researching English as a lingua franca: a conceptual enquiry. *International Journal of Applied Linguistics*, 19(2).
- Finegan, E. (2015). *Language: Its structure and use*. Cengage Learning.
- Firth, S. (1992). Pronunciation syllabus design: A question of focus. *Teaching American English pronunciation*.
- Flege, J. E. (1984). The Detection of French Accent by American Listeners. *The Journal of the Acoustical Society of America*, 76(3), 692-707.
- Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *The Journal of the Acoustical Society of America*, 84(1), 70-79.
- 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*, 233-277.
- Flege, J. E. (2005). Evidence for plasticity in studies examining second language speech acquisition. In *Workshop held at ISCA Workshop on Plasticity in Speech Perception, London, UK*.
- Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America*, 91(1), 370-389.
- Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Effects of age of second-language learning on the production of English consonants. *Speech Communication*, 16, 1-26.
- Flege, J. E., Frieda, E. M., & Nozawa, T. (1997). Amount of native-language (L1) use affects the pronunciation of an L2. *Journal of phonetics*, 25(2), 169-186.
- Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second language learning. *Journal of Memory and Language*, vol.41, 1999b, 78-104.
- Flege, J. E., MacKay, I. R., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, 23(4), 567-598.
- Flege, J. E., & MacKay, I. R. A. (2011). What accounts for “age” effects on overall degree of foreign accent? *Achievements and perspectives in the acquisition of second Language Speech: New Sounds*, 2, ed. M. Wremble et al., Frankfurt am Main: Peter Lang, 65-82.
- Flege, J. E., & Bohn, O. S. (2021). The Revised Speech Learning Model (SLM-r). *Second Language Speech Learning*, 1.
- Floare Bora, S. (2021). Taking literature off page! The effectiveness of a blended drama approach for enhancing L2 oral accuracy, pronunciation and complexity. *Language Teaching Research*, 1-12.
- Foot, J. A., Holtby, A. K., & Derwing, T. M. (2011). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada journal*, 1-22.
- Foot, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, 3(1), 34-56.
- Fouz-González, J. (2019). Podcast-based pronunciation training: Enhancing FL learners’ perception and production of fossilised segmental features. *ReCALL*, 31(2), 150-169.
- Fraser, H. (2001). *Teaching pronunciation: a guide for teachers of English as a second language*. Melbourne.
- Fraser, H. (2006). Helping teachers help students with pronunciation: A cognitive approach. *Prospect: A Journal of Australian TESOL*, 21(1), 78-94.

- Freeborn, L., & Rogers, J. (2019). Nonlinguistic factors that affect the degree of foreign accent in second language Mandarin. *Studies in Chinese Linguistics*, 40(1), 75-99.
- Frigols, M. J., & Marsh, D. (2007). CLIL as a catalyst for change in languages education. *Babylonia*, 3(07), 33-37.
- Fudge, E. (2015). *English Word-Stress*. Routledge.
- Galante, A. & Thomson, R.I. (2017), The Effectiveness of Drama as an Instructional Approach for the Development of Second Language Oral Fluency, Comprehensibility, and Accentedness. *TESOL Q*, 51, 115-142.
- Gatbonton E. Trofimovich P. & Magid M. (2005). Learners' ethnic group affiliation and L2 pronunciation accuracy: a sociolinguistic investigation. *Tesol Quarterly: A Journal for Teachers of English to Speakers of Other Languages and of Standard English As a Second Dialect* 489–512.
- Gatbonton, E., Trofimovich, P., & Segalowitz, N. (2011). Ethnic Group Affiliation and Patterns of Development of a Phonological Variable. *The Modern Language Journal*, 95, 188 - 204.
- Georgiou G.P. (2021) How Do Speakers of a Language with a Transparent Orthographic System Perceive the L2 Vowels of a Language with an Opaque Orthographic System? An Analysis through a Battery of Behavioral Tests. *Languages*, 6(3):118.
- Gholamrezapour, E., & Rabori, M. (2019). Empathy as a Predictor of Pronunciation Mastery: The Case of Female Iranian EFL Learners' Pronunciation Errors. *Journal of Teaching Language Skills*, 38(2), 107-129.
- Ghyasi, M., Yazdani, M., & Farsani, M. A. (2013). The relationship between personality types and self-regulated learning strategies of language learners. *International Journal of Applied Linguistics and English Literature*, 2(4), 74-82.
- Gilbert, J. B. (2010). Pronunciation as orphan: What can be done. *Speak Out*, 43(3), 6-10.
- Gilbert, J. B. (2012). *Clear speech teacher's resource and assessment book: Pronunciation and listening comprehension in North American English*. Cambridge University Press.
- Gilbert, J. B. (2014). Intonation is hard to teach. In Grant, L., & Brinton, D., *Pronunciation myths: Applying second language research to classroom teaching*, 107-137.
- Giles, H., Taylor, D. M., & Bourhis, R. (1973). Towards a theory of interpersonal accommodation through language: some Canadian data1. *Language in society*, 2(2), 177-192.
- Giles, H., Coupland, N., & Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. *Contexts of accommodation: Developments in applied sociolinguistics*, 1, 1-68.
- Giles, H., Powesland, P. (1997). Accommodation Theory. In: Coupland, N., Jaworski, A. (eds) *Sociolinguistics*. Modern Linguistics Series. Palgrave, London.
- Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (vol. 7, 7-28). Tilburg: Tilburg University Press.
- Golestani, N., Molko, N., Dehaene, S., LeBihan, D., & Pallier, C. (2007). Brain structure predicts the learning of foreign speech sounds. *Cerebral cortex*, 17(3), 575-582.
- Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain Language*, 109(2-3), 55-67.
- Gooch, R., Saito, K., & Lyster, R. (2016). Effects of recasts and prompts on L2 pronunciation development: Teaching English/ɪ/ to Korean adult EFL learners. *System*, 60, 117-127.

- Gordon, E. (1989). *Advanced measures of music audiation*. Gia Publications.
- Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. *Pronunciation in Second Language Learning and Teaching Proceedings*, 4(1), 194-206.
- 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.
- Gordon, J., & Darcy, I. (2019). Teaching segmentals vs. suprasegmentals: different effects of explicit instruction on comprehensibility. *Pronunciation in Second Language Learning and Teaching Proceedings*, 10(1), 116-126.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading*. Routledge, 34-57
- Granena, G., & Long, M. H. (2013). Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second language research*, 29(3), 311-343.
- Grant, L. (2001). *Well said: Pronunciation for clear communication* (2nd ed.). Heinle & Heinle.
- Grant, L. (2014). Prologue to the myths: What teachers need to know. *Pronunciation myths: Applying second language research to classroom teaching*, 1-33.
- Green, D. W. (1988). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67-81.
- Grivol, M. A., & Hage, S. R. D. V. (2011). Phonological working memory: a comparative study between different age groups. *Jornal da Sociedade Brasileira de Fonoaudiologia*, 23, 245-251.
- Grzega, J. (2005). Reflections on Concepts of English for Europe. *Journal for EuroLinguistiX* 2, 44-64.
- Guiora, A. Z., Beit-Hallahmi, B., Brannon, R. C. L., Dull, C. Y., & Scovel, T. (1972). The effects of experimentally induced changes in ego states on pronunciation ability in a second language: An exploratory study. *Comprehensive Psychiatry*, 13(5), 421-428.
- Guiora, A. Z., Brannon, R. C. L., & Dull, C. Y. (1972). Empathy and Second Language Learning. *Language Learning*, 22, 111-130.
- Hahn, L. D. (2004). Primary Stress and Intelligibility: Research to Motivate the Teaching of Suprasegmentals. *TESOL Quarterly*, 38(2), 201-223.
- Hanani, A., Russell, M. J., & Carey, M. J. (2013). Human and computer recognition of regional accents and ethnic groups from British English speech. *Computer Speech & Language*, 27(1), 59-74.
- Hancock, M. (2003). *English Pronunciation in Use: self-study and classroom use*. Cambridge Cambridge University Press.
- Hancock, M. (2007). *English Pronunciation in Use Elementary with Answers*. Cambridge University Press.
- Hancock, M. (2017). *English Pronunciation in Use Intermediate with Answers*, 2nd Edition. Cambridge University Press.
- Hancock, M. (2018). *Authentic Listening Resource Pack*. Delta Publishing.
- Hansen, J. G. (2004). Developmental sequences in the acquisition of English L2 syllable codas. *Studies in Second Language Acquisition*, 26, 85-124.

- Hardison, D. M. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24(4), 495-522.
- Harrison, G. (2012). "Oh, you've got such a strong accent": Language identity intersecting with professional identity in the human services in Australia. *International Migration*, 51, 192-204.
- Hayes-Harb, R., Brown, K., & Smith, B. L. (2018). Orthographic input and the acquisition of German final devoicing by native speakers of English. *Language and speech*, 61(4), 547-564.
- Henrichsen, L. (2019, March). A System for Analyzing and Evaluating Computer-Assisted Second-Language Pronunciation-Teaching Websites and Mobile Apps. In *Society for Information Technology & Teacher Education International Conference*, 963-968.
- Herd, W., Jongman, A., & Sereno, J. (2013). Perceptual and production training of intervocalic /d, r, r/in American English learners of Spanish. *The Journal of the Acoustical Society of America*, 133(6), 4247-4255.
- Herschensohn, J. R. (2007). *Language Development and Age*. Cambridge University Press.
- Hinton, M. (2012). An aptitude for speech: The importance of mimicry ability in foreign language pronunciation. In *Teaching and researching English accents in native and non-native speakers* (103-111). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Hinton, M. (2014). Aptitude as a cognitive/affective construct and its role in the language classroom. *Konińskie Studia Językowe*, 2(2), 125-141.
- Hirst, D. (1998). Intonation in British English. *Intonation systems. A survey of twenty languages*, 56-77.
- Hişmanoğlu, M. (2012). An investigation of pronunciation learning strategies of advanced EFL learners. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 43.
- Honeybone, P. & Watson, K. (2006). *Phonemes, Graphemes and Phonics for Liverpool English*.
- Honikman, B. (1964). *Articulatory Settings*. Longman, London, 73-84.
- Hu, X., & Reiterer, S. M. (2009). Personality and pronunciation talent in second language acquisition. *Language Talent and Brain Activity*, 97-129.
- Hu, X., Ackermann, H., Martin, J. A., Erb, M., Winkler, S., & Reiterer, S. M. (2013). Language aptitude for pronunciation in advanced second language (L2) learners: Behavioural predictors and neural substrates. *Brain and language*, 127(3), 366-376.
- Huang, Y. (2014). *Pragmatics*. Oxford University Press, USA.
- Hulstijn, J. H. (2015). Explaining phenomena of first and second language acquisition with the constructs of implicit and explicit learning: The virtues and pitfalls of a two-system view. In P. Rebuschat (ed.), *Implicit and Explicit Learning of Languages*, Amsterdam: John Benjamins Publishing, 25-46.
- Hyltenstam, K., & Abrahamsson, N. (2000). Who can become native-like in a second language? All, some, or none? On the maturational constraints controversy in second language acquisition. *Studia linguistica*, 54(2), 150-166.
- Hymes, D. (1972). On communicative competence. *sociolinguistics*, 269-293.
- Ioup, G., Boustagui, E., El Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in second language acquisition*, 16(1), 73-98.

- Isaacs, T. (2009). Integrating form and meaning in L2 pronunciation instruction. *TESL Canada Journal*, 27(1), 1-12.
- Isaacs, T., & Trofimovich, P. (2011). Phonological memory, attention control, and musical ability: Effects of individual differences on rater judgments of second language speech. *Applied Psycholinguistics*, 32(1), 113-140.
- Isaacs, T., & Thomson, R. (2013). Rater Experience, Rating Scale Length, and Judgments of L2 Pronunciation: Revisiting Research Conventions. *Language Assessment Quarterly*, 10, 135-159.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford University Press.
- Jenkins, J. (2007). *English as a Lingua Franca: Attitude and Identity*. Oxford University Press.
- Jenkins, J. (2012). English as a Lingua Franca from the classroom to the classroom. *ELT Journal*, 66(4), 486-494.
- Jesney, K. (2004). The use of global foreign accent rating in studies of L2 acquisition. *Calgary, AB: University of Calgary Language Research Centre Reports*, 1-44.
- Jilka, M. (2000). Testing the contribution of prosody to the perception of foreign accent. *New Sounds*, 4, 199-207.
- John, P., and Cardoso, W. (2009). Francophone ESL learners' difficulties with English /h/. In B Baptista, A Rauber, & M Watkins (eds.), *Recent Research in Second Language Phonetics/Phonology: Perception and Production*, Newcastle Upon Tyne: Cambridge Scholars, 118-140.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive psychology*, 21(1), 60-99.
- Jolliffe, D., & Farrington, D. P. (2006). Examining the relationship between low empathy and bullying. *Aggressive Behavior: Official Journal of the International Society for Research on Aggression*, 32(6), 540-550.
- Jones, D. (1926). *An English Pronouncing Dictionary; showing the pronunciation of over 50,000 words in international phonetic transcription* (Rev.). J.M. Dent; E.P. Dutton.
- Jones, R. H. (1997). Beyond "listen and repeat": Pronunciation teaching materials and theories of second language acquisition. *System*, 25(1), 103-112.
- Jones, R. H., & Evans, S. (1995). Teaching pronunciation through voice quality. *ELT Journal*, 49(3), 244-251.
- Jułkowska, I. A., & Cebrian, J. (2015). Effects of listener factors and stimulus properties on the intelligibility, comprehensibility and accentedness of L2 speech. *Journal of Second Language Pronunciation*, 1(2), 211-237.
- Kachlicka, M., Saito, K., & Tierney, A. (2019). Successful second language learning is tied to robust domain-general auditory processing and stable neural representation of sound. *Brain and Language*, 192, 15-24.
- Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: the English language in the Outer Circle. *English in the world: teaching and learning the language and the literatures*, R. Quirk & H. Widdowson (eds.), Cambridge: Cambridge University Press in association with The British Council, 11-30.
- Kachru, B. B. (ed.). (1992). *The other tongue: English across cultures*. Urbana: University of

- Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of experimental psychology: General*, 132(1), 47.
- Kang, O. (2010). Relative salience of suprasegmental features on judgments of L2 comprehensibility and accentedness. *System*, 38(2), 301-315.
- Kang, O., & Rubin, D. (2009). Reverse Linguistic Stereotyping: Measuring the Effect of Listener Expectations on Speech Evaluation. *Journal of Language and Social Psychology*, 28, 441-456.
- Kassaian, Z. (2011). Age and Gender Effect in Phonetic Perception and Production. *Journal of Language Teaching and Research*, 2.
- Katz, L., & Feldman, L. B. (1983). Relation between pronunciation and recognition of printed words in deep and shallow orthographies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9(1), 157.
- Kelly, L. G. (1969). *25 Centuries of Language Teaching*. Newbury House.
- Kendrick, H. (1997). Keep them talking! A project for improving students' L2 pronunciation. *System*, 25(4), 545-560.
- Kennedy, S., & Trofimovich, P. (2008). Intelligibility, Comprehensibility, and Accentedness of L2 Speech: The Role of Listener Experience and Semantic Context. *The Canadian Modern Language Review / La revue canadienne des langues vivantes*, 64, 459 - 489.
- Kennedy, S., & Trofimovich, P. (2010). Language awareness and second language pronunciation: A classroom study. *Language awareness*, 19(3), 171-185.
- Kennedy, S., & Trofimovich, P. (2017). Pronunciation acquisition. In S. Loewen and M. Sato (eds.) *The Routledge Handbook of Instructed Second Language Acquisition*, 260-279.
- Kim, T. (2008). Accentedness, comprehensibility, intelligibility, and interpretability of NNESTs. *The CATESOL Journal*, 20(1), 7-26.
- Kim, D. D., & Margolis, D. P. (1999). Teaching English pronunciation to Koreans: Development of an English pronunciation test EPT. In *KOTESOL proceedings: The Second Pan Asian Conference*, 89-112.
- Kinsella C. & Singleton D. (2014). Much more than age. *Applied Linguistics* 441–462.
- Kirby, J., & Giang, Đ. L. (2021). Relating production and perception of L2 tone. *Second language speech learning*, 249-272.
- Kissling, E. M. (2013). Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners?. *The modern language journal*, 97(3), 720-744.
- Korkut, P., & Çelik, Ö. (2021). Developing pronunciation through creative drama. *The Language Learning Journal*, 49(2), 147-159.
- Kormos, J., & Sáfár, A. (2008). Phonological short-term memory, working memory and foreign language performance in intensive language learning. *Bilingualism: Language and cognition*, 11(2), 261-271.
- Kouritzin, S. G. (1999). *Face [t] s of first language loss*. Routledge.
- Kralova, Z., Nemcokova, K., & Birova, J. (2020). Contrastive vs non-contrastive meta-phonetic input in teaching foreign language pronunciation. *Lidil. Revue de linguistique et de didactique des langues*, (61).

- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon Press.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Kraut, R., & Wulff, S. (2013). Foreign-accented speech perception ratings: A multifactorial case study. *Journal of Multilingual and Multicultural Development*, 34(3), 249-263.
- Kremmel, B., Eberharter, K., Konrad, E., Guggenbichler, E., Moser-Frötscher, D., Ebner, V., & Spöttl, C. (2023). The CEFR Companion Volume: Opportunities and challenges for language assessment. *Didaktik slawischer Sprachen*, 65-83.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence leads to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134.
- Labov, W. (1972). *Sociolinguistic Patterns*. Philadelphia, PA: University of Pennsylvania Press.
- Lambert, S. (1988). Shadowing. *Meta: Translators' Journal*, 37(2), 263-273.
- Lappin-Fortin, K., & Rye, B. J. (2014). The use of pre-/posttest and self-assessment tools in a French pronunciation course. *Foreign Language Annals*, 47(2), 300-320.
- LaScotte, D., & Tarone, E. (2019). Heteroglossia and constructed dialogue in SLA. *The Modern Language Journal*, 103, 95-112.
- LaScotte, D., Meyers, C., & Tarone, E. (2021). Voice and mirroring in SLA: Top-down pedagogy for L2 pronunciation instruction. *RELC Journal*, 52(1), 144-154.
- Lascotte, D. K., & Tarone, E. (2022). Channeling Voices to Improve L2 English Intelligibility. *The Modern Language Journal*, 106(4), 744-763.
- Laver, J. (1980). The phonetic description of voice quality. *Cambridge Studies in Linguistics London*, 31, 1-186.
- Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36(3), 345-366.
- Lee, J. W., & Yoon, K. O. (2019). Alternative Vocabulary Learning Approaches in EFL Setting: Bottom-Up or Top-Down?. *English Teaching*, 74(3), 141-160.
- Lefkowitz, N., & Hedgcock, J. (2002). Sound Barriers: Influences of Social Prestige, Peer Pressure and Teacher (Dis)Approval on FL Oral Performance. *Language Teaching Research*, 6(3), 223-244.
- Lenneberg, E. H. (1967). *Biological Foundations of Language*. New York: Wiley.
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093-1096.
- Lev-Ari, S., & Peperkamp, S. (2013). Low inhibitory skill leads to non-native perception and production in bilinguals' native language. *Journal of Phonetics*, 41(5), 320-331.
- LeVelle, K., & Levis, J.M. (2014). Understanding the impact of social factors on L2 pronunciation: Insights from learners. In J Levis & A Moyer (eds.), *Social dynamics in second language accent*, 97-118.
- Levis, J. M. (2005). Changing Contexts and Shifting Paradigms in Pronunciation Teaching. *TESOL Quarterly*, 39(3), 369-376.
- Levis, J. M. (2017). Evidence-based pronunciation teaching: A pedagogy for the future. *Journal of Second Language Pronunciation*, 3(1), 1-8.

- Levis, J. M., & McCrocklin, S. (2018). Reflective and effective teaching of pronunciation. In *Issues in applying SLA theories toward reflective and effective teaching*, 77-89.
- Levis, J. M., & Wu, A. (2018). Pronunciation Research into Practice and Practice Into Research. *CATESOL Journal*, 30(1), 1-12.
- Levis, J.M. (2020). Revisiting the Intelligibility and Nativeness Principles. *Journal of Second Language Pronunciation*, 6(3), 310-328.
- Levy, B. J., McVeigh, N. D., Marful, A., & Anderson, M. C. (2007). Inhibiting your native language: The role of retrieval-induced forgetting during second-language acquisition. *Psychological Science*, 18(1), 29-34.
- Lewandowski, N., & Jilka, M. (2019). Phonetic convergence, language talent, personality and attention. *Frontiers in Communication*, 4, 18.
- Lewis, C., & Deterding, D. (2018). Word Stress and Pronunciation Teaching in English as a Lingua Franca Contexts. *CATESOL Journal*, 30(1), 161-176.
- Ley, D. (nd.) 3 minutes to a proper British accent with University of Alberta Prof. David Ley. *University of Alberta Alumni Magazine*. <https://youtu.be/ldAKIzq7bvs>.
- Li, S. (2016). The construct validity of language aptitude: A meta-analysis. *Studies in second language acquisition*, 38(4), 801-842.
- Lim, S. J., & Holt, L. L. (2011). Learning foreign sounds in an alien world: Videogame training improves non-native speech categorization. *Cognitive science*, 35(7), 1390-1405.
- Lindhout P, Ale B (2009), "Language problems, an underestimated danger in major hazard control?", *Journal of Hazardous Materials* 172, 247–255.
- Lindsey, G. (2019). *English after RP: Standard British pronunciation today*. Springer.
- Lippi-Green, R. (2012). *English with an accent: Language, ideology, and discrimination in the United States*. New York: Routledge.
- Loewen, S. (2011). Focus on form. In *Handbook of research in second language teaching and learning*, 576-592.
- Long, M. H. (1990). Maturation constraints on language development. *Studies in Second Language Acquisition*, 12, 251-285.
- Long, M. H. (2000). Focus on form in task-based language teaching. *Language policy and pedagogy: Essays in honor of A. Ronald Walton*, 179-192.
- Lorenzen, J. (2019). Tonal Recall: Musical Ability and Toneme Recognition. *Pronunciation in Second Language Learning and Teaching Proceedings*, 10(1).
- Luo, B. (2016). Evaluating a computer-assisted pronunciation training (CAPT) technique for efficient classroom instruction. *Computer assisted language learning*, 29(3), 451-476.
- Luo, D., Shimomura, N., Minematsu, N., Yamauchi, Y., & Hirose, K. (2008). Automatic pronunciation evaluation of language learners' utterances generated through shadowing. In *Ninth annual conference of the international speech communication association*, 2807-2810.
- Lybeck, K. (2002). Cultural Identification and Second Language Pronunciation of Americans in Norway. *The Modern Language Journal*, 86, 174 – 191.
- MacDonald, S. (2018). Pronunciation Tutorials: Not Only Sounds, but Also Awareness of Self and Context. *CATESOL Journal*, 30(1), 249-265.

- Major, R. C. (2007). Identifying a Foreign Accent in an Unfamiliar Language. *Studies in Second Language Acquisition*, 29(4), 539–556.
- Major, R. C., Fitzmaurice, S. F., Bunta, F., & Balasubramanian, C. (2002). The effects of nonnative accents on listening comprehension: Implications for ESL assessment. *TESOL quarterly*, 36(2), 173–190.
- Mandell, J. (n.d.) Tonedead Test : Mandell, J. (n.d.) Retrieved from: www.jakemandell.com/tonedead
- Martinsen, R., Montgomery, C., & Willardson, V. (2017). The effectiveness of video-based shadowing and tracking pronunciation exercises for foreign language learners. *Foreign Language Annals*, 50(4), 661–680.
- McDermott, W.L.C. (1986). *The Scalability of Degrees of Foreign Accent*. Ph.D. Dissertation. Cornell University, Ithaca, NY.
- McGregor, A., & Reed, M. (2018). Integrating pronunciation into the English language curriculum: A framework for teachers. *CATESOL Journal*, 30(1), 69–94.
- Meara, P. (2005). LLAMA language aptitude tests: The manual. *Swansea: Lognostics*.
- Meritan, C., & Mroz, A. (2019). Impact of self-reflection and awareness-raising on novice French learners' pronunciation. *Foreign Language Annals*, 52(4), 798–821.
- Meyers, C. (2016). The Mirroring Project: A contextualized and integrated approach to pronunciation teaching. Presentation given at the Center for Advanced Research on Second Language Acquisition, University of Minnesota.
- Meyers, C. (2018). Mirroring: A Top-Down Approach to Pronunciation Instruction. *Pronunciationforteachers.com / Teaching Techniques*
- Miller, Carol. (2011). Auditory Processing Theories of Language Disorders: Past, Present, and Future. Language, speech, and hearing services in schools. 42.
- Miller, L., & Ng, R. (1994). Peer Assessment of Oral Language Proficiency. Perspectives: Working Papers of the Department of English, City Polytechnic of Hong Kong, 6, 41–56.
- Miller, S. (2001). Self-monitoring, self-help, and the route to intelligible speech. *The CATESOL Journal*, 13(1), 173–182.
- Milovanov, R. (2009). The connectivity of musical aptitude and foreign language learning skills: Neural and behavioural evidence. *Anglicana Turkuensia* 27, 1–56.
- Milovanov, R., & Tervaniemi, M. (2011). The interplay between musical and linguistic aptitudes: a review. *Frontiers in psychology*, 2, 321.
- Missaglia, F. (1999). Contrastive prosody in SLA: An empirical study with adult Italian learners of German. In *Proceedings from: 14th International Congress of Phonetic Sciences*, 551–554.
- Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. *Foreign language learning: Psycholinguistic studies on training and retention*, 339–364.
- Mompean, J. A. (2007). Estuary English: Revisiting the debate on its status as a new accent of English and potential EFL pronunciation model. *Bells: Barcelona English language and literature studies*.
- Mora, J. C., & Darcy, I. (2017). The Relationship Between Cognitive Control and Pronunciation in a Second Language. *Second Language Pronunciation Assessment: Interdisciplinary Perspectives*, (Eds.) T. Isaacs & P. Trofimovich, 95– 120.

- Mora, J. C., & Mora-Plaza, I. (2019). Contributions of cognitive attention control to L2 speech learning. *A sound approach to language matters—In honor of Ocke-Schwen Bohn*, 477-499.
- Mora, J. C., & Darcy, I. (2023). Individual differences in attention control and the processing of phonological contrasts in a second language. *Phonetica*, 80(3-4), 153-184.
- Moreno, L. (2016). *Channeling Charlie: Suprasegmental pronunciation in a second language learner's performance of others' voices*. MA Thesis, University of Minnesota.
- Morley, J. (1996). Georgetown University Round Table on Languages and Linguistics (GURT) 1996, *Linguistics, Language Acquisition, and Language Variation: Current Trends and Future Prospects*, 140.
- Moskovsky, C., Jiang, G., Libert, A., & Fagan, S. (2015). Bottom-up or top-down: English as a foreign language vocabulary instruction for Chinese university students. *Tesol Quarterly*, 49(2), 256-277.
- 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. (2004). *Age, accent, and experience in second language acquisition: an integrated approach to critical period inquiry* (Vol. 7). Multilingual Matters.
- Moyer, A. (2008). Conceptions of L2 phonology: Integrating cognitive and sociolinguistic approaches to research and teaching.
- Moyer, A. (2009). Input as a critical means to an end: Quantity and quality of experience in L2 phonological attainment. *Input matters in SLA*, 159, 174.
- Moyer, A. (2010). Do gender differences in L2 accent really exist. In *Meeting of American Association for Applied Linguistics (AAAL)*, Atlanta, GA.
- Moyer, A. (2011). An Investigation of Experience in L2 Phonology: Does Quality Matter More than Quantity? *Canadian Modern Language Review-revue Canadienne Des Langues Vivantes*, 67, 191-216.
- Moyer, A. (2013). *Foreign accent: The phenomenon of non-native speech*. Cambridge University Press.
- Moyer, A. (2014a). Exceptional Outcomes in L2 Phonology: The Critical Factors of Learner Engagement and Self-Regulation, *Applied Linguistics*, 35 (4), 418–440.
- Moyer, A. (2014b). What's age got to do with it? Accounting for individual factors in second language accent. *Studies in Second Language Learning and Teaching*, 4(3), 443-464.
- Moyer, A. (2015a). *Autonomy in second language phonology: Choice vs. limits*. Sixth Annual Pronunciation in Second Language learning and Teaching Conference, 1-17.
- Moyer, A. (2015b). *Foreign Accent: The Phenomenon of Non-Native Speech*. Cambridge University Press.
- Moyer, A. (2016). The puzzle of gender effects in L2 phonology. *Journal of Second Language Pronunciation*, 2, 8-28.
- Moyer, A. (2017). Autonomy in second language phonology: Choice vs. limits¹. *Language Teaching*, 50(3), 395-411.
- Moyer, A. (2018). An Advantage for Age? Self-Concept and Self-Regulation as Teachable Foundations in Second Language Accent. *CATESOL Journal*, 30(1), 95-112.
- Moyer, A. (2021). *The Gifted Language Learner: A case of nature or nurture?*. Cambridge University Press.

- Muller, G., & Levis, J. (2016). Integrating pronunciation into listening/speaking classes. *Integrating pronunciation with other language skills*, 27-42.
- Müller, J. L., Friederici, A. D., & Männel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Sciences*, 109(39), 15953-15958.
- Muñoz, C. (2014). Starting age and other influential factors: Insights from learner interviews. *Studies in Second Language Learning and Teaching*, 4(3), 465-484.
- Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44(1), 1-35.
- Munro, M. J. (1995). Nonsegmental factors in foreign accent: Ratings of filtered speech. *Studies in Second Language Acquisition*, 17(1), 17-34.
- Munro, M. J. (2018). How well can we predict second language learners' pronunciation difficulties? *The CATESOL Journal*, 30(1), 267- 282.
- Munro, M. J., & Derwing, T. M. (1994). Evaluations of Foreign Accent in Extemporaneous and Read Material. *Language Testing*, 11(3), 253-266.
- Munro, M. J., & Derwing, T. M. (1995a). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73-97.
- Munro, M. J., & Derwing, T. M. (1995b). Processing Time, Accent, and Comprehensibility in the Perception of Native and Foreign-Accented Speech. *Language and Speech* 38(3), 289-306.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech the role of speaking rate. *Studies in second language acquisition*, 23(4), 451-468.
- Munro, M. J., Derwing, T. M., & Burgess, C. S. (2003). The detection of foreign accent in backwards speech. In M J Sole, D Recasens & Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona: Futurgraphic, 535-538.
- Munro, M. J., Derwing, T. M. & Sato, K. (2006). Salient accents, covert attitudes: consciousness-raising for pre-service second language teachers. *Prospect: an Australian journal of TESOL*, 21(1), 67-79.
- Murphy, J. (2014). Teacher training programs provide adequate preparation in how to teach pronunciation. In Linda J. Grant & Donna Brinton (eds.) *Pronunciation Myths: Applying Second Language Research to Classroom Teaching* (, pp. 188-224), Michigan Press., 188-224.
- Myers, C. (2016). The Mirroring Project: A contextualized and integrated approach to pronunciation teaching. Presentation given at the Center for Advanced Research on Second Language Acquisition, University of Minnesota.
- 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.
- Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, 44(2), 580-605.
- Nardo, D., & Reiterer, S. (2009). Musicality and phonetic language aptitude. *Language talent and brain activity*, 213-256.

- Nu, D. T. (2015). Application of Contrastive Analysis in teaching English pronunciation to Vietnamese adults. *Prague, Czech Republic May 21-23, 2015*, 30-33.
- O'Brien, M. G. (2014). L2 learners' assessments of accentedness, fluency, and comprehensibility of native and nonnative German speech. *Language Learning*, 64(4), 715-748.
- O'Brien, M. G. (2021). Ease and difficulty in L2 pronunciation teaching: A mini-review. *Frontiers in Communication*, 5, 1-7.
- Olson, L. L., & Samuels, J. (1973). The relationship between age and accuracy of foreign language pronunciation. *The Journal of Educational Research*, 66(6), 263-268.
- Orii-Akita, M. (2014). The effectiveness of interactive teaching methods in EFL classrooms: a comparison with bottom-up and top-down methods. In *Proceedings of the International Symposium on the Acquisition of Second Language Speech, Concordia working Papers in Applied Linguistics*, 5, 464-477.
- Oyama, S. (1976). A sensitive period for the acquisition of nonnative phonological system. *Journal of Psycholinguistic Research*, 5, 261-283.
- Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in psychology*, 2, 142.
- Penfield, W. & Roberts, L. (1959). *Speech and brain mechanisms*. New York: Atheneum Press.
- Pennington, M.C. (1988). Teaching Pronunciation from the Top Down. *University of Hawaii Working Papers in ESL*, 7(1), 203-227.
- Pennington M. C. (1989) Teaching pronunciation from the top down. *RELC Journal* 20(1): 20–38.
- Pennington, M. C. (1996). *Phonology in English Language Teaching: An International Approach*. Longman.
- Pennington M. C. (1997) Phonology in language teaching: essentials of theory and practice. In: Bardovi-Harlig K, Hartford B (eds) *Beyond Methods: Components of Second Language Teacher Education*. New York: McGraw-Hill, 67–87.
- Pennington, M. C. (2019a). 'Top-down' pronunciation teaching revisited. *RELC Journal*, 50(3), 371-385.
- Pennington, M. C. (2019b). 7 Best Practices for Teaching Pronunciation. *TESOL Connections: October 2019*.
- Pennington, M. C., & Richards, J. C. (1986). Pronunciation revisited. *TESOL quarterly*, 20(2), 207-225.
- Pennington, M. C., & Waxler, R. P. (2017). *Why reading books still matters: The power of literature in digital times*. Routledge.
- Perani, D. (2005). The neural basis of language talent in bilinguals. *Trends in Cognitive Sciences*, 9(5), 211-213.
- Pérez-Ramón, R., Lecumberri, M. L. G., & Cooke, M. (2022). Foreign accent strength and intelligibility at the segmental level. *Speech Communication*, 137, 70-76.
- Pica, T. (1995). The language educator at work in the learner-centered classroom: Communicate, decision-make, and remember to apply the (educational) linguistics. *Educational linguistics, crosscultural communication, and global interdependence*, 264.

- Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of sociolinguistics*, 6(2), 179-208.
- 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.
- Piske, T., Flege, J. E., MacKay, I. R., & Meador, D. (2002). The production of English vowels by fluent early and late Italian-English bilinguals. *Phonetica*, 59(1), 49-71.
- Plunkett, J. (2011). Cheryl Cole 'Axed' from US X Factor. *The Guardian*, Guardian News and Media, 26 May 2011. www.theguardian.com/culture/2011/may/26/cheryl-cole-us-x-factor.
- Polat, N., & Mahalingappa, L. J. (2010). Gender differences in identity and acculturation patterns and L2 accent attainment. *Journal of Language, Identity, and Education*, 9(1), 17-35.
- Polka, L. (1995). Linguistic influences in adult perception of non-native vowel contrasts. *The Journal of the Acoustical Society of America*, 97(2), 1286-1296.
- Prabhu, N. S. (1987). *Second language pedagogy* (Vol. 20). Oxford: Oxford university press.
- Probst, K., Ke, Y., & Eskenazi, M. (2002). Enhancing foreign language tutors—in search of the golden speaker. *Speech Communication*, 37(3-4), 161-173.
- Purcell, E. T., & Suter, R. W. (1980). Predictors of Pronunciation Accuracy: A Reexamination. *Language Learning*, 30, 271 – 287.
- Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive highvariability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning & Technology*, 22(1), 69–96.
- Rato, A. (2014). Effects of perceptual training on the identification of English vowels by native speakers of European Portuguese. In *Proceedings of the international symposium on the acquisition of second language speech*, 5, 529-546.
- Rato A., Rauber A. S., Soares L. P. & Lucas L. R. (2014). Challenges in the perception and production of English front vowels by native speakers of European Portuguese. *Revista Diacrítica*, 28 (1), 142–159.
- Rato, A. & Oliveira, D. (2022). Assessing the robustness of L2 perceptual training: A closer look at generalization and retention of learning (Chapter 13). In U. Alves & J Albuquerque (Eds.) *Second Language Pronunciation: Different Approaches to Teaching and Training*, 369-396.
- Reiterer, S. M., Hu, X., Erb, M., Rota, G., Nardo, D., Grodd, W., & Ackermann, H. (2011). Individual differences in audio-vocal speech imitation aptitude in late bilinguals: functional neuro-imaging and brain morphology. *Frontiers in psychology*, 2, 271.
- Reiterer, S. M., Hu, X., Sumathi, T. A., & Singh, N. C. (2013). Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Frontiers in psychology*, 4, 782.
- Ricard, E. (1986). Beyond Fossilization: A Course in Strategies and Techniques in Pronunciation for Advanced Adult Learners. *TESL Canada Journal*, 243-253.
- Richards, M., & Cotos, E. (2019). Segmental Accuracy: A Recommended Training Sequence for Moving Learners From Accurate Perception to Accurate and Automatic Production in the Stream of Speech. *Pronunciation in Second Language Learning and Teaching Proceedings*, 10(1).
- Riney, T. J., Takagi, N., & Inutsuka, K. (2005). Phonetic Parameters and Perceptual Judgments of Accent in English by American and Japanese Listeners. *TESOL Quarterly*, 39, 441-466.

- Robb, M. P., MacLagan, M. A., & Chen, Y. (2004) Speaking rates of American and New Zealand varieties of English. *Clinical Linguistics & Phonetics*, 18(1), 1-15.
- Rogerson-Revell, P. (2011). *English phonology and pronunciation teaching*. Bloomsbury Publishing.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of experimental psychology: General*, 104(3), 192.
- Rosch, E. (1978). Principles of categorization. In *Cognition and categorization* (pp. 27-48). Routledge.
- Rubin, D. (2012). The power of prejudice in accent perception: Reverse linguistic stereotyping and its impact on listener judgments and decisions. *Pronunciation in Second Language Learning and Teaching Proceedings*, 3(1).
- Rupp, L. (2018). *English Pronunciation in a Global World*. FutureLearn.
- Safronova, E., & Mora, J. C. (2012). Acoustic and phonological memory in L2 vowel perception. *Proceedings of 22nd EUROSLA*, 384-390.
- Safronova, E., & Mora, J. C. (2013). Attention control in L2 phonological acquisition. *Applied linguistics in the age of globalization*, 384-390.
- Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language awareness*, 20(1), 45-59.
- Saito, K. (2012). Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *Tesol Quarterly*, 46(4), 842-854.
- Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English/ɪ/to adult Japanese learners. *Language learning*, 63(3), 499-529.
- Saito, K. (2023). How does having a good ear promote successful second language speech acquisition in adulthood? Introducing auditory precision hypothesis-L2. *Language Teaching*, 56(4), 522-538.
- Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɪ/ by Japanese learners of English. *Language learning*, 62(2), 595-633.
- 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.
- Saito, K., Trofimovich, P., & Isaacs, T. (2017). Using Listener Judgments to Investigate Linguistic Influences on L2 Comprehensibility and Accentedness: A Validation and Generalization Study. *Applied Linguistics*, 38(4), 439-462.
- Saito, K., Dewaele, J. M., Abe, M., & 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.
- Saito, K., Kachlicka, M., Sun, H., & Tierney, A. (2020). Domain-general auditory processing as an anchor of post-pubertal second language pronunciation learning: Behavioural and neurophysiological investigations of perceptual acuity, age, experience, development, and attainment. *Journal of Memory and Language*, 115, 104168.
- Saito, K., Trofimovich, P., Abe, M., & In'nami, Y. (2020). Dunning-Kruger effect in second language speech learning: How does self-perception align with other perception over time? *Learning and Individual Differences*, 79, 1-10.

- Saito, K., Petrova, K., Suzukida, Y., Kachlicka, M., & Tierney, A. (2022). Training auditory processing promotes second language speech acquisition. *Journal of Experimental Psychology: Human Perception and Performance*, 48(12), 1410.
- Sakai, M., & Moorman, C. (2018). Can perception training improve the production of second language phonemes? A meta-analytic review of 25 years of perception training research. *Applied Psycholinguistics*, 39(1), 187-224.
- Sardegna, V. G. (2011). Pronunciation Learning Strategies That Improve ESL Learners' Linking. *Pronunciation in Second Language Learning and Teaching Proceedings*, 2(1), 105-121.
- Sardegna, V. G. (2022). Evidence in favor of a strategy-based model for English pronunciation instruction. *Language Teaching*, 55(3), 363-378.
- Sardegna, V. G., & McGregor, A. (2013). Scaffolding Students' Self-Regulated Efforts for Effective Pronunciation Practice. *Pronunciation in Second Language Learning and Teaching Proceedings*, 4(1), 182-193.
- Saritsas, B. B., & Saritas B. H. (2022) A comparison of Portuguese and English Phonology. *International Journal of Social Sciences and Management Review Volume 5* (2).
- Schmidt, R. W. (1990). The role of consciousness in second language learning¹. *Applied linguistics*, 11(2), 129-158.
- Schmidt, R. (2001). Attention. In P Robinson (Ed.), *Cognition and second language instruction*, 3-32.
- Schmidt, R. (2012). Attention, awareness, and individual differences in language learning. *Perspectives on individual characteristics and foreign language education*, 6, 27-49.
- Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. *Talking to learn: Conversation in second language acquisition*, 237, 326.
- Schmidtke, J. (2016). The Bilingual Disadvantage in Speech Understanding in Noise Is Likely a Frequency Effect Related to Reduced Language Exposure. *Frontiers in psychology*, 7, 678.
- Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language learning*, 28(1), 129-142.
- Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human speech*. Wadsworth Publishing Company.
- Sedlackova, K. (2007). *Some aspects of Non-native Acquisition of English Pronunciation in Portuguese Speakers*. (Doctoral dissertation, Masaryk University).
- Seidlhofer, B. (2005). English as a lingua franca. *ELT Journal* 59(4), 339–341.
- Seliger, H. W., Krashen, S.D., & Ladefoged, P. (1975). Maturational constraints in the acquisition of second language accent. *Language Sciences*, 36, 20–22.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10, 209-30.
- Sereno, J., Lammers, L., & Jongman, A. (2016). The relative contribution of segments and intonation to the perception of foreign-accented speech. *Applied Psycholinguistics*, 37(2), 303-322.
- Shao, Y., Saito, K., & Tierney, A. (2023). How does having a good ear promote instructed second language pronunciation development? Roles of domain-general auditory processing in choral repetition training. *TESOL Quarterly*, 57(1), 33-63.

- Shepherd, D. (2001). Portuguese Speakers. In Swan, M., & Smith, B. (Eds.), *Learner English: A teacher's guide to interference and other problems*, Cambridge University Press, 113-128.
- Showalter, C. E., & Hayes-Harb, R. (2015). Native English speakers learning Arabic: The influence of novel orthographic information on second language phonological acquisition. *Applied psycholinguistics*, 36(1), 23-42.
- Singleton, D. (2017). Language aptitude: Desirable trait or acquirable attribute?. *Studies in Second Language Learning and Teaching*, 7(1), 89-103.
- Singleton, D & Lesniewska, J. (2012). Age and SLA: Research highways and byeways. In M. Pawlak (ed.), *New perspectives on individual differences in language learning and teaching*, Heidelberg: Springer, 97-113.
- Skehan, P., Granena, G., Jackson, D. O., & Yilmaz, Y. (2016). Foreign language aptitude, acquisitional sequences, and psycholinguistic processes. *Cognitive individual differences in second language processing and acquisition*, 17-40.
- Slevc, L. R., & Miyake, A. (2006). Individual Differences in Second-Language Proficiency: Does Musical Ability Matter? *Psychological Science*, 17(8), 675–681.
- Smith, L. E., & Nelson, C. L. (1985). International intelligibility of English: directions and resources. *World Englishes*, 4, 333-342.
- Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and speech*, 20(4), 357-365.
- Sorace, A. (2003). Near-nativeness. *The handbook of second language acquisition*, 130-151.
- Sorace, A. (2011). Pinning down the concept of “interface” in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1-33.
- Sousa, E. M. A. de. (2017). *Production and perception of the English /h/: the case of native Portuguese speakers of English as a foreign language*. MA thesis.
- Southwood, M. H., & Flege, J. E. (1999). Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics and Phonetics*, 13(5), 335–349.
- Speech Accent Archive. (1999). George Mason University. <http://accent.gmu.edu/browse.php>
- Spezzini, S. (2004). English immersion in Paraguay: Individual and sociocultural dimensions of language learning and use. *International Journal of Bilingual Education and Bilingualism*, 7(5), 412-431.
- Stapp, Y. F., & Center, G. (1999). Neural plasticity and the issue of mimicry tasks in L2 pronunciation studies. *Neural Plasticity*, 3(4).
- Strange, W. (1995). Cross-language studies of speech perception: A historical review. *Speech perception and linguistic experience: Issues in cross-language research*, 3-45.
- Sturiale, M. (2002). RP: Received or Reference Pronunciation. *Linguistica e Filologia*, 15, 89–112.
- Suter, R. W. (1976). Predictors of Pronunciation Accuracy in Second Language Learning. *Language Learning*, 26(2), 233-253.
- Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *Canadian modern language review*, 50(1), 158-164.
- Swales, J. (1988). Discourse communities, genres and English as an international language. *World Englishes*, 7(2), 211-220.

- Szpyra-Kozłowska, J. (2014). *Pronunciation in EFL Instruction*. Channel View Publications.
- Tabandeh, F., Moinzadeh, A., & Barati, H. (2018). Tasks in Explicit L2 Pronunciation Instruction: FonF vs. FonFS in Improving Phonemic Accuracy and Comprehensibility. *3L: Southeast Asian Journal of English Language Studies*, 24(1), 112-127.
- Tahta, S., Wood, M., & Loewenthal, K. (1981). Foreign accents: Factors relating to transfer of accent from the first language to a second language. *Language and Speech*, 24(3), 265-272.
- Tarone, E., & Meyers, C. (2018). The Mirroring Project. *Speaking in a second language*, 17, 197-223.
- Thompson, I. (1991). Foreign accents revisited: The English pronunciation of Russian immigrants. *Language learning*, 41(2), 177-204.
- Thomson, R. I. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *Calico Journal*, 28(3), 744-765.
- Thomson, R. I. (2014). Accent reduction and pronunciation instruction are the same thing. *Pronunciation myths: Applying second language research to classroom teaching*, 160-187.
- Thomson, R. I. (2018a). Measurement of accentedness, intelligibility and comprehensibility. In O. Kang & A. Ginther (Eds.), *Assessment in second language pronunciation*, 11–29.
- Thomson, R. I. (2018b). High variability [pronunciation] training (HVPT) A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation*, 4(2), 208-231.
- Thomson, R. I. (2022). Perception in Pronunciation Training. *Second Language Pronunciation: Bridging the Gap Between Research and Teaching*, 42-60.
- Thomson, R. I., & Derwing, T. M. (2015a). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36(3), 326-344.
- Thomson, R. I. & Derwing, T. M. (2015b). Is Phonemic Training Using Nonsense or Real Words More Effective? *Pronunciation in Second Language Learning and Teaching Proceedings* 7, 88-97.
- Thornbury, S. (1993). Having a good jaw: Voice-setting phonology. *ELT Journal*, 47(2), 126-131.
- Tiewtrakul, T., & Fletcher, S. R. (2010). The challenge of regional accents for aviation English language proficiency standards: A study of difficulties in understanding in air traffic control–pilot communications. *Ergonomics*, 53(2), 229-239.
- Timmis, I. (2002). Native-Speaker Norms and International English: A Classroom View. *ELT Journal*, 56(3), 240-249.
- Todaka, Y. (1993). *A cross-language study of voice quality: Bilingual Japanese and American English speakers*. University of California, Los Angeles.
- Trofimovich, P., & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: Implications for pronunciation instruction. *The Modern Language Journal*, 90(4), 519-535.
- Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905-916.
- Trofimovich, P., Kennedy, S., & Foote, J. (2015). Variables affecting L2 pronunciation development. *The handbook of English pronunciation*, 353-373.
- Trofimovich, P., Isaacs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: Investigating self-and other-perception of second language speech. *Bilingualism: Language and Cognition*, 19, Cambridge University Press, 122-140.

- Trude, A. M., & Tokowicz, N. (2011). Negative transfer from Spanish and English to Portuguese pronunciation: The roles of inhibition and working memory. *Language Learning*, 61(1), 259-280.
- Trudgill, P., & Cheshire, J. (1989). Dialect and education in the United Kingdom. In J. Cheshire, V. Edwards, H. Münstermann & B. Weltens (eds.), *Dialect and Education: some European perspectives*. Clevedon: Multilingual Matters, 94-109.
- Trudgill, P., & Hannah, J. (2008). *International English: A guide to the varieties of Standard English (5th ed.)*. Routledge.
- Tsai, N., Au, J., & Jaeggi, S. M. (2016). Working memory, language processing, and implications of malleability for second language acquisition. *Cognitive individual differences in second language processing and acquisition*, 69-88.
- Tsunemoto, A., Trofimovich, P., Blanchet, J., Bertrand, J., & Kennedy, S. (2022). Effects of benchmarking and peer-assessment on French learners' self-assessments of accentedness, comprehensibility, and fluency. *Foreign Language Annals*, 55(1), 135-154.
- Uzal, M., Peltonen, T., Huotilainen, M., & Aaltonen, O. (2015). Degree of perceived accent in Finnish as a second language for Turkish children born in Finland. *Language Learning*, 65(3), 477-503.
- Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher–student interaction: A decade of research. *Educational psychology review*, 22, 271-296.
- Van Der Slik, F. W., Van Hout, R. W., & Schepens, J. J. (2015). The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PloS one*, 10(11), 1-22.
- Vangehuchten, L., Verhoeven, V., & Thys, P. (2015). Pronunciation proficiency and musical aptitude in Spanish as a foreign language: results of an experimental research project. *Revista de Lingüística y Lenguas Aplicadas*, 10, 90-100.
- Verbeke, G., & Simon, E. (2023). Listening to accents: Comprehensibility, accentedness and intelligibility of native and non-native English speech. *Lingua*, 292, 1-25.
- Veríssimo, J., Heyer, V., Jacob, G., & Clahsen, H. (2018). Selective effects of age of acquisition on morphological priming: Evidence for a sensitive period. *Language Acquisition*, 25(3), 315-326.
- 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.
- Vokic, G. (2011). When alphabets collide: Alphabetic first-language speakers' approach to speech production in an alphabetic second language. *Second Language Research*, 27(3), 391-417.
- Wach, A. (2011). Native Speaker and English as a Lingua Franca Pronunciation Norms: English Majors' view. *Studies in Second Language Learning and Teaching*, 1, 247-266.
- Walker, R. (2010). *Teaching the Pronunciation of English as a Lingua Franca*. Oxford: Oxford University Press.
- Wen, Z., Li, S., Schwieter, J., & Benati, A. (2019). Working memory in L2 learning and processing. *The Cambridge handbook of language learning*, 365-389.
- Wen, Z. E., & Skehan, P. (2021). Stages of Acquisition and the P/E Model of Working Memory: Complementary or contrasting approaches to foreign language aptitude?. *Annual Review of Applied Linguistics*, 41, 6-24.
- Wester, M., & Mayo, C. (2014). Accent rating by native and non-native listeners. *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing – Proceedings*, 7699-7703.

- Wheeler, P., & Saito, K. (2022). Second language speech intelligibility revisited: Differential roles of phonological accuracy, visual speech, and iconic gesture. *Modern Language Journal*, 106(2), 429–448.
- Wilkerson, M. E. (2010). Identifying Accent in German: A Comparison of Native and Non-Native Listeners. *Die Unterrichtspraxis*, 43, 144-153.
- Wilkerson, M. E. (2013). The sound of German: Descriptions of accent by native and non-native listeners. *Die Unterrichtspraxis*, 46(1), 106–118.
- Willis, J. (1996). A flexible framework for task-based learning. *Challenge and change in language teaching*, 52, 62.
- Wing, H. D. (1962). A revision of the Wing musical aptitude test. *Journal of Research in Music Education*, 10(1), 39-46.
- Wrembel, M. (2010). L2-accented speech in L3 production. *International Journal of Multilingualism*, 7(1), 75-90.
- Yang, Y., Shintani, N., Li, S., & Zhang, Y. (2017). The effectiveness of post-reading word-focused activities and their associations with working memory. *System*, 70, 38-49.
- Yeldham, M. (2016). Second language listening instruction: Comparing a strategies-based approach with an interactive, strategies/bottom-up skills approach. *Tesol Quarterly*, 50(2), 394-420.
- Yeni-Komshian, G. H., Flege, J. E., & Liu, S. (2000). Pronunciation proficiency in the first and second languages of Korean–English bilinguals. *Bilingualism: Language and cognition*, 3(2), 131-149.
- Yoshida, M. T. (2016). *Beyond Repeat after Me: Teaching Pronunciation to English Learners*. TESOL Press. Available from: TESOL International Association. 1925 Ballenger Avenue Suite 550, Alexandria, VA 22314.
- Yoshida, M. T. (2018). Choosing technology tools to meet pronunciation teaching and learning goals. *CATESOL Journal*, 30(1), 195-212.
- Young-Scholten, M. (2013). Great expectations in phonology? Second language acquisition research and its relation to the teaching of older and younger learners. In M Whong, K-H Gil & H Marsden (eds.), *Universal Grammar and the Second Language Classroom*, Dordrecht: Springer, 207-229.
- Zárate-Sández, G. (2017). Reexamining Foreign Accent: How Much Can Personality Explain? *Ilha do Desterro*, 70(3), 227-243.
- Zhang, R., & Yuan, Z. M. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition*, 42(4), 905-918.
- Zheng, C., Saito, K., & Tierney, A. (2022). Successful second language pronunciation learning is linked to domain-general auditory processing rather than music aptitude. *Second Language Research*, 38(3), 477-497.

APPENDICES

APPENDIX A - CEFR (CV) Descriptors

		A1	A2	B1	B2	C1	C2
U N D E R S T A N D I N G	Listening	I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly.	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local area, employment). I can catch the main point in short, clear, simple messages and announcements.	I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.	I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.	I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.	I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.
	Reading	I can understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues.	I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters.	I can understand texts that consist mainly of high frequency everyday or job-related language. I can understand the description of events, feelings and wishes in personal letters.	I can read articles and reports concerned with contemporary problems in which the writers adopt particular attitudes or viewpoints. I can understand contemporary literary prose.	I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.	I can read with ease virtually all forms of the written language, including abstract, structurally or linguistically complex texts such as manuals, specialised articles and literary works.
S P E A K I N G	Spoken Interaction	I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.	I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.
	Spoken Production	I can use simple phrases and sentences to describe where I live and people I know.	I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job.	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.	I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.	I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.
W R I T I N G	Writing	I can write a short, simple postcard, for example sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.	I can write short, simple notes and messages relating to matters in areas of immediate needs. I can write a very simple personal letter, for example thanking someone for something.	I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions.	I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.	I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select style appropriate to the reader in mind.	I can write clear, smoothly-flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.

Appendix B – Speech Token Instructions Email

Hello participant,

the pronunciation course will be starting in two weeks, so I would now like to ask you to record a speech sample. This will be used to compare with the speech samples you provide after the course to see if any change has been achieved. Please follow the instructions as follows:

- You are going to record a 2-MINUTE speech sample on your mobile phone. If you don't have recording software, the free Rev app is recommended as it works on iPhone or Android: <https://www.rev.com/voicerecorder>
- In the speech sample, I would like you to talk about the topic at the bottom of this email - **ONLY SCROLL DOWN TO THE TOPIC WHEN YOU ARE READY TO RECORD**
- The speech sample should not be prepared, just your natural speech, so don't write anything down or plan what you will say
- Make sure that you are in a quiet room with minimal background noise and that you will not be interrupted
- When you are ready, read the topic at the end of this email. Allow yourself no more than 10 seconds to think about it and then record your answer
- Send the sample to:
- Your samples will only be used anonymously for scientific purposes and will not be shared with third parties.
- I would be grateful if you could send the speech samples by Friday, 25th February 2022

Once I have received the speech sample, I will send you a Zoom link to the classes on your chosen day. As always, many thanks for your help with this project.

Accent Rating Questionnaire

Practice Exercise: Speaker 0

* Indicates required question

1. Name: *

READ through the questions below. Then CLICK ON LINK to listen and RATE THE SPEAKER according to the questions. Give your GENERAL IMPRESSION of the speaker, not a detailed analysis. SPEAKER 0 is a PRACTICE EXERCISE, so you may want to listen MORE THAN ONCE.

https://drive.google.com/file/d/1YW_mE37R_Q9x7N7KHZHGVzgnQmHJyl_96/view?usp=sharing

2. Rate the speaker: how close to a NATIVE SPEAKER does the accent sound? *

Mark only one oval.

1 2 3 4 5 6 7 8 9

Not ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Native speaker

How would you rate the following:

2a. PRONUNCIATION OF INDIVIDUAL SOUNDS/WORDS (eg. how 'th' is pronounced; how words are stressed; ...)

Mark only one oval.

1 2 3 4 5 6 7 8 9

Not ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Native speaker

2b. SENTENCE LEVEL FEATURES (eg. the 'music' of the sentences ; the rhythm of the sentences ; ...)

Mark only one oval.

1 2 3 4 5 6 7 8 9

Not ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Native speaker

2c. Overall FLUENCY (eg. the speech is fluent; speaker is slow or hesitant; ...)

Mark only one oval.

1 2 3 4 5 6 7 8 9

Not ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Native speaker

2d. Overall VOCABULARY FEATURES (eg odd use of words; good range of words; ...)

Mark only one oval.

1 2 3 4 5 6 7 8 9

Not ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Native speaker

2e. Overall LANGUAGE ACCURACY (eg. words in wrong order; no grammar mistakes; ...)

Mark only one oval.

1 2 3 4 5 6 7 8 9

Not ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Native speaker

(Optional) GIVE FURTHER DETAILS about your decisions above (eg. sounds like an Italian accent; 'th' sound pronounced as 'd')

3. Which NATIVE SPEAKER ENGLISH ACCENT does the speaker most closely approximate? *

Mark only one oval.

- ☐ British English accent
- ☐ American English accent
- ☐ No native speaker accent
- ☐ I don't know
- ☐ Other: _____

Thank you!

This content is neither created nor endorsed by Google.

Google Forms

Learner Consent Form and Pronunciation Questionnaire

* Indicates required question

Consent Form

WHAT IS THE GOAL?

This is an investigation into ways of working on pronunciation in the language classroom.

WHO CAN PARTICIPATE?

Preference will be given to Portuguese native speakers who have an upper-intermediate level of English or above and who started learning English after the age of four.

TASKS

As a participant, you will be asked to complete the following tasks:

1. To fill in this short form about your pronunciation profile (2 min).
2. To record a sample of your spoken English before the course, at the end of the course and 3 months after the end of the course (about 1-2 min each).
3. To attend a series of 10 60-minute classes working towards achieving more native-like, standard British English pronunciation.

GUARANTEES TO PARTICIPANTS

1. Your participation will be voluntary.
2. You may withdraw from the study at any time, without any consequences.
3. Your data will only be used for scientific purposes and will not be shared with third parties.
4. You may ask for your data to be removed from the study at any time in the future.
5. All information you provide will be treated anonymously.
6. You will not suffer any emotional, physical, psychological or other damage as a result of this research.
7. You will receive feedback on your current accent and a certificate of participation in the study on completion of the course.

OTHER INFORMATION:

If you have any questions about this study, please contact: juliemason@fcsh.unl.pt

Thank you for your interest in participating in this study!

1. I consent to the use of the information in this questionnaire and the voice recordings associated with the course for research purposes and understand that my data will be treated anonymously at all times. *

Mark only one oval.

- ☐ Yes
☐ No

2. Name *

3. Gender *

Mark only one oval.

- ☐ Female
☐ Male
☐ Other: _____

4. Age *

5. Email *

6. Which course do you wish to attend? *

Mark only one oval.

- ☐ Fridays 6-7pm
☐ Saturdays 10-11am
☐ I don't mind

7. The 10 sessions will run from the 4/5 March to the 13/14 May with a one-week break for Easter week. I confirm that I will be able to attend all the sessions:

Mark only one oval.

- ☐ Yes
☐ No
☐ Other: _____

8. Will you be taking any other courses which involve work on pronunciation during the period of this course?

Mark only one oval.

- ☐ Yes
☐ No
☐ Other: _____

PRONUNCIATION QUESTIONNAIRE

1. What is/are your nationality/nationalities? (You may tick more than one box) *

Tick all that apply.

- ☐ Portuguese
☐ Brazilian
☐ Other: _____

2. What is/are your mother tongue(s)? (You may tick more than one box) *

Tick all that apply.

- ☐ European Portuguese
☐ Brazilian Portuguese
☐ Other: _____

3. What is your current level of English? *

Mark only one oval.

- ☐ B1 (Intermediate)
☐ B2 (Upper-intermediate)
☐ C1 (Advanced)
☐ C2 (Upper-advanced/near-native speaker)
☐ Unknown

What proof do you have of your current level eg. a certificate or school grade? Please explain below: *

4. At what age were you first exposed to English? *

Mark only one oval.

- ☐ Prior to the age of 5
- ☐ 5-6 years
- ☐ 7-11 years
- ☐ 10-17 years
- ☐ 18 years +
- ☐ Other: _____

5. What did the initial exposure involve? (You may tick more than one box) *

Tick all that apply.

- ☐ Watching content (eg. TV programmes/films) **with** subtitles
- ☐ Watching content (eg. TV programmes/films) **without** subtitles
- ☐ Listening to spoken English
- ☐ Reading in English
- ☐ Speaking in English
- ☐ Other: _____

6. At what age did you start having English lessons? *

Mark only one oval.

- ☐ Prior to the age of 5
- ☐ 5-6 years
- ☐ 7-11 years
- ☐ 10-17 years
- ☐ 18 years +
- ☐ Other: _____

7. How would you describe your current accent in English? *

Mark only one oval.

- ☐ British English accent
- ☐ American English accent
- ☐ Another English native speaker accent
- ☐ Portuguese accent in English
- ☐ I don't know what I sound like in English
- ☐ No accent
- ☐ Other: _____

8. Do you have difficulty with any particular sounds or features of spoken English? (You may tick more than one box) *

Tick all that apply.

- ☐ Individual sounds (eg. th sound; h at the beginning of a word; ...)
- ☐ Word stress (ie. where to accent a longer word)
- ☐ Sentence level features (eg. the 'music' or rhythm of sentences)
- ☐ Slow or hesitant speech
- ☐ Other: _____

9. (Optional) Add more details about the above question if necessary

10. What would be the accent which you would like to have in English? *

Mark only one oval.

- ☐ Standard British English accent
- ☐ Standard American English accent
- ☐ I'm happy with my current accent
- ☐ Other: _____

Thank you for your answers!

This content is neither created nor endorsed by Google.

Google Forms

End of Course Feedback

Please take a few minutes to provide some feedback on the course which you attended

* Indicates required question

I consent to the use of the information in this questionnaire for research purposes and understand that *
my data will be treated anonymously at all times.

Mark only one oval.

☐ Yes

☐ No

Name *

Email *

1. Do you feel the course has given you an INCREASED AWARENESS OF YOUR ACCENT in English? *

Tick all that apply.

☐ Yes

☐ No

2. WHICH FEATURES OF YOUR ACCENT do you have an increased awareness of? (you may tick more *
than one box)

Tick all that apply.

☐ Individual sounds (eg. th sound; h at the beginning of a word; ...)

☐ Word stress (eg. where to accent a longer word)

☐ Sentence level features (eg. the 'music' or rhythm of sentences)

☐ Fluency (eg. slow or hesitant speech)

☐ Other:

3. What would be the ACCENT WHICH YOU WOULD LIKE TO HAVE in English? *

Mark only one oval.

- ☐ Standard British English accent
- ☐ Standard American English accent
- ☐ I'm happy with my current accent
- ☐ Other: _____

4. Do you feel that the course has HELPED YOU TO CHANGE your pronunciation? *

Tick all that apply.

- ☐ Yes
- ☐ No
- ☐ I'm happy with my current accent
- ☐ Other: _____

5. If the course helped you to change, WHAT ASPECTS OF YOUR PRONUNCIATION has it helped you to change? (you may tick more than one box)

Tick all that apply.

- ☐ Individual sounds (eg. th sound; h at the beginning of a word; ...)
- ☐ Word stress (eg. where to accent a longer word)
- ☐ Sentence level features (eg. the 'music' or rhythm of sentences)
- ☐ Fluency (eg. speed of speech; hesitancy)
- ☐ Other: _____

6. WHICH ACTIVITIES WERE MOST USEFUL in helping you to work on pronunciation? (you may tick more than one box) *

Tick all that apply.

- ☐ Work on listening to the target accent
- ☐ Work on individual sounds (eg. /th/ sound; live-leave distinction; ...)
- ☐ Work on suprasegmental features (eg. word stress; sentence rhythm; music of speech)
- ☐ Work on comparing British and American English
- ☐ Work on evaluating your current accent
- ☐ Work on factors influencing accent (eg. child-adult differences; accent and identity)
- ☐ Work on global strategies to change pronunciation (eg. copying a target speaker; roleplay)
- ☐ Other: _____

Please add any OTHER COMMENTS you have about the course:

Many thanks for your participation and for your continued help with this research!

This content is neither created nor endorsed by Google.

Google Forms