Training L2 Speech Segmentation with Word-Spotting

Janise Farrell

Thesis submitted for the degree of Doctor of Philosophy
University of Western Sydney
2015
Dedicated to the memory of my mother,

Wendy Ann Shanley

who filled me with a love of books and learning

and reminded me I was strong.
Acknowledgements

I would first like to express my sincere thanks to my primary supervisor, Professor Anne Cutler, for her tireless support, encouragement and mentorship. Thank you, Anne, for teaching me how to think and write like a psycholinguist. I truly appreciate the support you have given me on both a professional and personal level. I would also like to thank my co-supervisors, Professor Denis Burnham and Associate Professor Jeesun Kim. Denis, I am very grateful for your feedback and advice during the data analysis stage of Experiments Three and Four. Thank you, Jeesun, for first recommending me for the project and providing tremendous support in establishing it.

I would also like to thank Dr Natalie Morrison for her advice when commencing the project and Johnson Chen for automating the search of the Macquarie Dictionary, programming software to translate the Macquarie code version of the phonetic script into IPA and frequently saving the day. I am also very grateful to Dr Michael Tyler for creating the artificial language learning (ALL) materials used in Experiment Five. Special thanks also to Dr Bronson Harry for helping me record the dialogues used in the connected speech processes training and Dr Erin Cvejic for teaching me so much about DMDX and SPSS. Many thanks to both of you for all the laughs in those early years.

I would especially like to thank everyone at UWSCollege, particularly Margaret Miller for her support in allowing me to conduct so many experiments at the college,
Cammy Chan for administrative support with a smile, and Joanne Nimmons for being so encouraging and lovely to talk to. I’m very grateful to Chris Duckworth, former technical officer at UWSCollege, for patiently trouble-shooting the many technical issues I experienced while running experiments at the college. I am also particularly grateful to the many ESL students at UWSCollege and the University of Western Sydney, who acted as participants in this study, and also to the many hundreds of ESL students I have taught, and who have taught me, throughout my career. My efforts were, as always, for you.

I would also like to express my appreciation to Sonya O’Shanna for liaison and support through difficult times, and to Darlene Williams and Colin Schoknecht for all you do for everyone, every day.

Last, but certainly not least, I am thankful to my husband, Stephen Farrell, for his love, support, patience, proofreading, pep-talks and cooked meals. Stephen, I truly could not have achieved this without you. Thank you from the bottom of my heart.
The work presented in this thesis is, to the best of my knowledge and belief, original except as acknowledged in the text. I hereby declare that I have not submitted this material either in full or in part, for a degree at this or any other institution. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project’s design and conception is acknowledged.
Table of Contents

List of Tables.............................................................................................................. xiv
List of Figures.............................................................................................................. xvi
Abstract.................................................................................................................. xviii

1 Introduction........................................................................................................... 1
  1.1. The Research Question............................................................................. 1
  1.2. The Aims of the Research...................................................................... 5
  1.3. Outline of the Chapters.......................................................................... 7

2 Speech Recognition and Segmentation......................................................... 8
  2.1. Speech Recognition.............................................................................. 8
  2.2. Segmentation Cues............................................................................. 10
    2.2.1. Lexicality..................................................................................... 11
    2.2.2. Statistical Cues........................................................................... 12
    2.2.3. Allophones................................................................................. 13
    2.2.4. Prosodic Cues............................................................................. 14
      2.2.4.1. Rhythm............................................................................... 14
      Stress............................................................................................... 14
      Syllables......................................................................................... 16
      Morae.............................................................................................. 17
    2.2.4.2. Intonation.............................................................................. 19
    2.2.5. Phonotactic Cues....................................................................... 20
  2.3 Second Language Speech Segmentation................................................. 21
2.3.1 Non-L1 Listening
2.3.2 Bilingual Listening
2.3.3 Second Language Listening
2.4 Learning Second Language Segmentation Strategies
2.4.1 Prosody
2.4.1.1 Suppression of Rhythmic Cues
2.4.1.2 Acquisition of Prosodic Cues
2.4.2 Phonotactics
2.5 Summary
2.6 Can Second Language Segmentation be Trained?

3 Training L2 Speech Segmentation
3.1 Introduction
3.2 Background
3.3 Decoding Approaches
3.3.1 Cauldwell's Spontaneous Speech Model
3.3.2 Field's Process Approach
3.3.2.1 The Process Approach to Teaching Segmentation
3.4 The Metacognitive Approach
3.4.1 The Metacognitive Approach to Teaching Segmentation
3.5 Top-Down or Bottom-Up?
3.6 The Training of Decoding Processes
3.6.1 Dictation
3.6.1.1 Computer Assisted Dictation Training
3.6.2 Explicit Training in Connected Speech Processes
5.3.3.1 Pre-test................................................................. 105
5.3.3.2 Word-Spotting.................................................... 105
5.3.3.3 Dictation............................................................. 106
5.3.3.4 Connected Speech Processes............................... 107
5.3.3.5 Post-tests and Questionnaire............................... 108
5.3.3.6 Number of Participants Completing Each Session..... 108

5.4 Results................................................................................. 109
  5.4.1 Pre-test............................................................................ 109
  5.4.2 Pre-test – Post-test 1...................................................... 110
  5.4.3 Pre-test – Post-test 2...................................................... 111
  5.4.4 Pre-test – Post-test 3...................................................... 112
  5.4.5 Pre-test and All Post-tests............................................. 113
  5.4.6 Questionnaire............................................................... 114

5.5 Discussion.......................................................................... 116

6 Experiment Two: Pilot Test.................................................. 118
  6.1 Aim.................................................................................. 118
  6.2 Background and Rationale for the New Tests.................. 118
    6.2.1 Listening Cloze Test................................................... 119
    6.2.2 Partial Dictations: Dialogues................................. 122
  6.3 Method............................................................................ 124
    6.3.1 Participants.............................................................. 124
    6.3.2 Stimuli.................................................................... 124
      6.3.2.1 Listening Cloze Test........................................... 124
        Sentences.................................................................. 124
        Design of Bleep Sentences........................................ 125
        The Selection of Words to Replace with Bleeps............ 126
6.3.2.2 Partial Dictations ................................................. 127
6.3.3 Procedure .................................................................................. 128
  6.3.3.1 Pre-test .............................................................................. 128
  6.3.3.2 Training Programs .......................................................... 129
     Word-Spotting ........................................................................ 129
     Connected Speech Processes .................................................. 129
6.4 Cessation of Pilot Test ................................................................. 130
6.5 Results .......................................................................................... 130
  6.5.1 Listening Cloze Test ............................................................ 130
  6.5.2 Partial Dictations ................................................................. 130
     6.5.2.1 Notes on the Marking Scheme .................................... 130
     6.5.2.2 Partial Dictation Results ............................................. 132
     6.5.2.3 Missegmentations ....................................................... 133
     Correlations ............................................................................ 133
  6.5.3 Correlations between Listening Cloze Test and Dictations .... 133
6.6 Discussion ..................................................................................... 133
  6.6.1 Listening Cloze Test ............................................................ 133
  6.6.2 Partial Dictation ................................................................. 135
6.7 Future Directions ........................................................................ 136
7 Experiment Three ........................................................................... 137
  7.1 Aim, Design and Hypotheses .................................................. 137
  7.2 Method ...................................................................................... 138
     7.2.1 Participants .................................................................... 138
     7.2.2 Stimuli ............................................................................ 140
        7.2.2.1 Training Programs .................................................. 140
        7.2.2.2 Pre- and Post-tests .................................................. 140
Dictations................................................................. 140

Mini-Dictations.......................................................... 141

7.2.3 Equipment.................................................................. 142

7.2.4 Procedure.................................................................. 142

7.2.4.1 Pre-test................................................................. 144

7.2.4.2 Training Programs.................................................. 144

Word-Spotting............................................................... 144

Connected Speech Processes............................................. 145

7.2.4.3 Post-test Instruments............................................. 145

Daily Mini-Dictations....................................................... 145

Post-tests......................................................................... 145

7.3 Results......................................................................... 145

7.3.1 Dictations................................................................... 146

7.3.1.1 Pre-test................................................................. 146

Missegmentations.......................................................... 146

Correlations................................................................. 146

7.3.1.2 Pre-test – Post-test 1.............................................. 146

Missegmentations.......................................................... 148

Correlations................................................................. 149

7.3.1.3 Pre-test – Post-test 2.............................................. 149

Missegmentations.......................................................... 151

Correlations................................................................. 152

7.3.2 Mini-Dictations......................................................... 152

7.3.2.1 Session 1............................................................. 152

Missegmentations.......................................................... 153

Correlations................................................................. 154
7.3.2.2 Session 2................................................................. 154
  
  Missegmentations.................................................. 155
  Correlations......................................................... 156

7.3.3 Word-Spotting Scores................................................. 156
  7.3.3.1 Correlations..................................................... 157
  7.3.4 Questionnaire...................................................... 157

7.4 Discussion........................................................................ 160

8 Experiment Four.................................................................. 162

  8.1 Aim, Design and Hypotheses................................................. 162
  8.2 Method........................................................................ 163
    8.2.1 Participants.......................................................... 163
    8.2.2 Stimuli.................................................................... 165
      8.2.2.1 Training Programs............................................ 165
      8.2.2.2 Pre- and Post-tests........................................... 165
      8.2.2.3 Reaction Time Post-test..................................... 166
        Vocabulary Teaching............................................... 167
        Post-test................................................................ 169
  8.2.3 Equipment............................................................... 169
  8.2.4 Procedure.................................................................. 169

8.3 Results............................................................................ 170
  8.3.1 Dictations................................................................. 170
    8.3.1.1 Pre-test............................................................ 170
      Missegmentations.................................................... 171
      Correlations......................................................... 172
    8.3.1.2 Pre-test – Post-test 1......................................... 172
      Missegmentations.................................................... 174
9.2.3 Equipment.................................................................................. 202
9.2.4 Procedure.................................................................................. 202
9.2.4.1 Daily Training Schedule...................................................... 202
9.2.4.2 ALL Post-test........................................................................ 203
9.3 Results.......................................................................................... 204
9.3.1 ALL Post-test............................................................................ 204
9.3.1.1 Exploring First Language Groups............................... 208
    Indo-Aryan Listeners............................................................ 209
9.3.2 Word-Spotting Scores............................................................... 212
9.3.2.1 Mean Percentage Correct.................................................. 212
9.3.2.2 Mean Percentage Segmented Correctly...................... 213
9.3.3 Correlations............................................................................ 214
9.4 Discussion..................................................................................... 214
10 General Discussion......................................................................... 216
10.1 Summary.................................................................................... 216
10.2 Evaluation.................................................................................... 217
10.2.1 Viability.................................................................................. 217
10.2.2 Efficacy.................................................................................. 218
10.3 Limitations of the Study............................................................... 220
10.3.1 Low Participant Numbers...................................................... 220
10.3.2 Pre- and Post-Tests................................................................. 221
10.3.3 Mini-Dictations...................................................................... 221
10.3.4 Design of Materials............................................................... 222
10.4 Recommendations...................................................................... 222
10.4.1 Pedagogical Implications....................................................... 222
10.4.1.1 Training............................................................................. 222
Participants’ Reflections..................................................224

10.4.1.2 Testing........................................................................224

Listenig Cloze Test.................................................................225

Dictation.............................................................................225

Missegmentations.................................................................226

10.4.2 Recommendations for Further Research.......................226

10.5 Conclusion.....................................................................229

References............................................................................230
# Appendices

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Words Excluded From Automated Search</td>
<td>250</td>
</tr>
<tr>
<td>B</td>
<td>Full list of Embedded Word-Spotting Training Stimuli</td>
<td>269</td>
</tr>
<tr>
<td>C</td>
<td>The IELTS 9-Band Scale</td>
<td>287</td>
</tr>
<tr>
<td>D</td>
<td>Pilot Study: Questionnaire</td>
<td>289</td>
</tr>
<tr>
<td>E</td>
<td>Experiment One: Dictation Texts</td>
<td>290</td>
</tr>
<tr>
<td>F</td>
<td>Experiment One: Questionnaire</td>
<td>299</td>
</tr>
<tr>
<td>G</td>
<td>Experiment One: Results</td>
<td>303</td>
</tr>
<tr>
<td>H</td>
<td>Experiment Two: Listening Cloze Pre-test Sentence List</td>
<td>305</td>
</tr>
<tr>
<td>I</td>
<td>Experiment Two: Partial Dictation</td>
<td>308</td>
</tr>
<tr>
<td>J</td>
<td>Experiment Three: Dictation Sentences</td>
<td>310</td>
</tr>
<tr>
<td>K</td>
<td>Experiment Three: Mini-Dictation Sentences</td>
<td>316</td>
</tr>
<tr>
<td>L</td>
<td>Experiment Three: Results</td>
<td>326</td>
</tr>
<tr>
<td>M</td>
<td>Experiments Three &amp; Four: Questionnaire</td>
<td>332</td>
</tr>
<tr>
<td>N</td>
<td>Experiment Four: Vocabulary Teaching</td>
<td>335</td>
</tr>
<tr>
<td></td>
<td>Day One</td>
<td>335</td>
</tr>
<tr>
<td></td>
<td>Day Two</td>
<td>338</td>
</tr>
<tr>
<td>O</td>
<td>Experiment Four: Results</td>
<td>342</td>
</tr>
<tr>
<td>P</td>
<td>Experiment Five: Connected Speech Processes Training</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>Day One</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>Day Two</td>
<td>355</td>
</tr>
<tr>
<td></td>
<td>Day Three</td>
<td>359</td>
</tr>
<tr>
<td>Q</td>
<td>Experiment Five: Results</td>
<td>363</td>
</tr>
</tbody>
</table>
List of Tables

5.1 Experiment One: The composition of the three training groups........... 96
5.2 Experiment One: Training schedule............................................. 105
5.3 Experiment One: Completion rates............................................. 108
5.4 Experiment One: Questionnaire results....................................... 115
7.1 Experiment Three: The composition of the two training groups...... 138
7.2 Experiment Three: Completion rates............................................. 139
7.3 Experiment Three: Training schedule........................................... 142
7.4 Experiment Three: Questionnaire results....................................... 158
8.1 Experiment Four: The composition of the two training groups........ 164
8.2 Experiment Four: Completion rates............................................. 164
8.3 Experiment Four: Words used in the reaction time post-test......... 166
8.4 Experiment Four: Training schedule............................................ 170
8.5 Experiment Four: Mean pre-test dictation scores........................... 171
8.6 Experiment Four: Mean pre-test missegmentation scores.............. 172
8.7 Experiment Four: Correlations - partial mini-dictation and
missegmentation scores.................................................................. 179
8.8 Experiment Four: Questionnaire results........................................ 187
9.1 Experiment Five: Composition of the Training x Exposure Groups.... 195
9.2 Experiment Five: Language backgrounds...................................... 196
9.3 Experiment Five: Mean participant age ....................................... 196
9.4 Experiment Five: Mean length of residence................................... 197
9.5 Experiment Five: The ‘words’ of each artificial language............... 199
9.6 Experiment Five: Mean artificial language learning task scores........ 205
9.7 Experiment Five: Mean score of Language x ALL Exposure groups.... 207
9.8 Experiment Five: Mean scores of each first language group within
the L2 Group............................................................................. 208

O.1 Experiment Four: Significant pairwise comparisons
- partial mini-dictations............................................................... 314

O.2 Experiment Four: Significant pairwise comparisons
- whole-dictations....................................................................... 315
List of Figures

4.1 Validation: Mean correct word detection ...................................................... 89
4.2 Validation: Questionnaire results ................................................................. 90
5.1 Experiment One: Individual pre-test scores .................................................. 109
5.2 Experiment One: Mean pre-test – post-test 1 scores ..................................... 111
5.3 Experiment One: Mean pre-test – post-test 2 scores ..................................... 112
5.4 Experiment One: Mean pre-test – post-test 3 scores ..................................... 113
5.5 Experiment One: Mean scores across pre-test and all post-tests .............. 114
5.6 Experiment One: Helpfulness ratings for each training program .............. 114
6.1 Experiment Two: Frequency distribution of partial dictation scores .......... 132
7.1 Experiment Three: Mean pre-test - post-test 1 dictation scores .............. 147
7.2 Experiment Three: Mean pre-test - post-test 1 missegmentations ......... 149
7.3 Experiment Three: Mean pre-test – post-test 2 dictation scores .............. 150
7.4 Experiment Three: Mean pre-test - post-test 2 missegmentations ........... 151
7.5 Experiment Three: Mean session 1 mini-dictation scores ....................... 153
7.6 Experiment Three: Mean session 1 mini-dictation missegmentations ........... 154
7.7 Experiment Three: Mean session 2 mini-dictation scores ....................... 155
7.8 Experiment Three: Mean session 2 mini-dictation Missegmentations ........... 156
7.9 Experiment Three: Mean correct word-spotting responses ................... 157
7.10 Experiment Three: Questionnaire results ................................................ 159
8.1 Experiment Four: Mean pre-test – post-test 1 dictation scores................. 173
8.2 Experiment Four: Mean pre-test - post-test 1 missegmentations............. 174
8.3 Experiment Four: Mean pre-test – post-test 2 dictation scores............. 176
8.4 Experiment Four: Mean pre-test - post-test 1 missegmentations............. 176
8.5 Experiment Four: Mean session 1 partial mini-dictation scores.............. 178
8.6 Experiment Four: Mean session 1 partial mini-dictation
missegmentations ................................................................................. 178
8.7 Experiment Four: Mean session 1 whole mini-dictation scores.............. 180
8.8 Experiment Four: Mean session 1 whole mini-dictation
Missegmentations .................................................................................. 181
8.9 Experiment Four: Mean correct word-spotting responses..................... 183
8.10 Experiment Four: Mean correct word-spotting responses by
training group ............................................................................................. 184
8.11 Experiment Four: Mean percentage of correctly segmented
word-spotting items .................................................................................. 185
8.12 Experiment Four: Questionnaire results.............................................. 189
9.1 Experiment Five: Mean ALL task scores by language group............... 205
9.2 Experiment Five: Mean correct word-spotting responses by
language group ............................................................................................ 212
9.3 Experiment Five: Mean percentage of correctly segmented items by
language group ............................................................................................ 213
Abstract

The segmentation of continuous speech is a difficult task for second language learners. Listeners develop language-specific strategies to find word boundaries, based on the phonological structure of their first language. Second language listeners apply the strategies appropriate for their L1 when listening to their L2, resulting in processing delays and inaccuracy. However, research shows that listeners can learn segmentation strategies appropriate for their L2 and suppress those of their L1. This study explored whether the segmentation skills appropriate for English could be trained implicitly using the word-spotting task. The performance of L2 listeners trained with the experimental method was compared to those who received established training methods: dictation (in Experiment One) and training in connected speech processes (in Experiments One to Five). Group results were compared using a test of listening comprehension (in Experiment One), dictation (in Experiment Three), both dictation and a test of reaction times in detecting newly taught vocabulary (in Experiment Four) and an artificial language learning (ALL) task (in Experiment Five). A listening cloze test was trialled in Experiment Two. The inter-group difference in improvement was insignificant in each experiment, but the group receiving word-spotting training displayed more improvement in most measures throughout the study. We conclude therefore that the implicit, computer-based word-spotting technique 1) matches dictation and teacher-led, metalinguistic training in connected-speech processes in affecting listening competence, 2) is a viable method of training L2 segmentation skills and 3) is worthy of further investigation with larger participant numbers.
Chapter 1

Introduction

1.1 The Research Question

Listening to speech seems such an effortless and automatic process that we underestimate its complexity; that is, unless one happens to be a language learner. Listening to a second language is a difficult task. One of the challenges involved is speech segmentation: the ability to recognise individual words in a stream of continuous speech. Listeners take the apparent ease with which they segment their native language for granted, but a foreign language will often sound like one long, unbroken stream of syllables – and that is exactly what speech is. Under the conventions of written language, spaces are provided between words to guide the reader as to where they start and end. Unfortunately, there is no direct equivalent to these spaces in speech. Words are not separated by pauses.

In order to comprehend speech, listeners need to be able to segment its continuous stream into individual words using acoustic cues. Varying sublexical cues have been proposed and include distributional probabilities (Saffran, Aslin, & Newport, 1996a; Saffran, Newport & Aslin, 1996b; van der Lught, 2001), context-sensitive allophones (Altenberg, 2005; Jusczyk, Hohne & Bauman, 1999a), phonotactic (speech sound sequence) constraints (Mattys, Jusczyk, Luce & Morgan, 1999; McQueen 1998;
Vroomen, Tuomainen & de Gelder, 1998) and prosodic cues (Cutler & Norris, 1988; Jusczyk, Houston and Newsome, 1999b; Mehler, Dommergues, Frauenfelder & Segui, 1981; Vroomen, et al., 1998). Lexical theories (Mattys, White & Melhorn, 2005, White, Melhorn & Mattys, 2010), including the Possible Word Constraint (PWC), a speech segmentation strategy based on the assessment of whether posited word boundaries will create viable residues (Norris, McQueen, Cutler & Butterfield, 1997) have also been proposed. Listeners make use of multiple cues in speech segmentation (see for example, Johnson & Jusczyk, 2001; Mattys et al., 2005).

Speech segmentation strategies are language-specific. As will be discussed in depth in Chapter 2, the cues employed are part of the speech processing competence of the native listener, which is shaped by the phonological structure of his or her native language (see for instance, Aquil, 2012; Cutler, Mehler, Norris, & Segui, 1986; Cutler & Norris, 1988; Kim, Davis & Cutler, 2008; McQueen, 1998; Murty, Otake, & Cutler, 2007; Otake, Hatano, Cutler, & Mehler, 1993; Tyler & Cutler, 2009; Vroomen, et al., 1998).

The prosodic cue to segmentation relies on the rhythmic structure of the listener’s native language. English has stress-based rhythm (Abercrombie, 1967; Pike, 1945) and English speakers use the predominance of initial stress (perceptually realised in English in the suprasegmental parameters of pitch, duration and amplitude) to find word onsets (Cutler & Norris, 1988). In contrast, French has syllable-based rhythm (Abercrombie, 1967) and French listeners segment speech based on syllable rhythm (Mehler, et al., 1981). Native English listeners also apply implicit knowledge about the phonotactic sequences allowed in English (for example, /br/) and disallowed (for example, /dn/) to locate word boundaries (McQueen, 1998).
One factor in the difficulty experienced in non-native language listening is that listeners attempt to use their native language speech segmentation strategies when they listen to a second language. For instance, French listeners have been shown to segment English by syllabic structure (Cutler, et al., 1986), and it has been demonstrated that German listeners place word boundaries between the sequences /sl/ and /tw/, which, while not permissible sequences in German, are perfectly permissible onsets in English (Weber & Cutler, 2006).

The use of first language segmentation strategies in a second language is counterproductive. English learning listeners are more likely than native English listeners to arrive at incorrect segmentations and more reluctant to discard these despite further evidence to the contrary (Field, 2008c). Altenberg (2005) found that Spanish learners of English segmented speech correctly (on the basis of allophonic variation absent in Spanish) only 76% of the time. In addition, Goh (2000) found that 8 out of 10 of the self-reported listening problems reported in the diaries of Chinese English language learners could be classified as being related to speech segmentation and perceptual processing.

The difficulty of second language speech segmentation impacts on listeners’ performance and causes them anxiety. Graham (2006) reported that speech segmentation was one of the most significant listening problems described by English learners of French. The learners also described listening as their weakest language skill. Bekleyen (2009) found that the failure to recognise 'sentence segments' was a major factor in the high levels of anxiety experienced by trainee English teachers in Turkey when listening to English. In addition, in his review of second language listening comprehension research, Vandergrift (2007) described speech segmentation as a major problem for second language listeners.
As will be discussed in depth in Chapter 2, the good news is that some L2 segmentation strategies may be acquired and some inappropriate L1 strategies inhibited (Hanulíková, Mitterer & McQueen, 2011; Weber, 2000; Weber & Cutler, 2006, and see Cutler (2001) for a discussion). In particular, it seems with enough linguistic experience, non-native listeners are able to exploit L2 phonotactic cues and to suppress L1 rhythmic segmentation strategies. It seems that native-like listening becomes possible with higher levels of exposure and ability.

This potential has been recognised in the second language acquisition (SLA) literature (see for instance, Field, 2008c; Vandergrift, 2007). Yet, despite the fact that the majority of foreign language teachers surveyed by Graham, Santos and Francis-Brophy (2014) agreed that learners’ main listening problems lie in segmentation difficulties, most also reported never asking students to complete activities that would strengthen segmentation skills. Part of the problem, for teachers at all levels of experience, might be confusion as to how this may be done. In their study, Graham et al. (2014) listed ‘identify word boundaries’ as a learner activity. This is a vague sounding enterprise, rather than a concrete way to train the identification of words in continuous speech. However, it is important to also note that 67% of the teachers surveyed by Graham et al. reported never asking students to ‘transcribe’ (or, complete dictation activities), which is a simple way word boundaries may have been considered in the classroom and students’ success in finding them assessed.

Clearly, it is important that attention is given within the field of second language teaching research to the development of speech segmentation training methodology. As will be discussed in Chapter 3, some suggestions have been made. However, despite the importance of speech segmentation skills and the promise that they may
be taught, little attention has been given to the evaluation of relevant training methodology.

Only one previous study has investigated the effects of training English language learners in the speech segmentation strategies applied by native listeners. Al-jasser (2008) explicitly trained participants in the phonotactic cues used in native English segmentation. This study will be considered in depth in Chapters 2 and 3. It is important here to note that the study was limited to explicit, teacher-centred, classroom-based instruction and failed to provide ample exposure to the rhythmic structure of English. No studies to date have trained L2 segmentation skills implicitly or incorporated student-led technology.

1.2 The Aims of the Research

This thesis describes a research program that aimed to evaluate the efficacy of a novel method of training second language speech segmentation strategies. The word-spotting method provides implicit exposure to the phonotactic and rhythmic cues used by native listeners to segment English. Repeated exposure to these cues has the potential to provide sufficient linguistic experience for L2 listeners to start to inhibit inappropriate L1, and apply appropriate L2, segmentation strategies.

The word-spotting task was first developed by Cutler and Norris (1988) and has hitherto been used in psycholinguistic research as a tool for studying speech segmentation (see McQueen, 1996). It was adapted here to function as a training program with the aim of creating an improvement in the segmentation of continuous speech, and hence, word recognition and listening comprehension. In the task, listeners are asked to identify real words embedded in nonsense strings. For instance, the stimulus item feskept contains the nonsense syllable fes followed by the target
word *kept*. When used in psycholinguistic experiments, word-spotting materials usually feature some filler items which do not contain embedded target words. However, in the use of the task here as a training method, every item contains an embedded word. Listeners type their response and then receive feedback on the correct answer before moving to the next item. The training was designed to provide English learners with repeated exposure to the cues used by native English speakers to segment speech.

Several studies discussed above employed the word-spotting task (including Cutler and Norris, 1988; Hanulíková et al., 2011; McQueen, 1998; van der Lught, 2001; Weber and Cutler, 2006). Cutler and Shanley (2010) conducted a validation test of the task and found that performing the task was achievable and enjoyable for the target group.

To test the efficacy of the word-spotting training method, training outcomes were compared with those of groups receiving established methods of training: dictation (in Experiment One) and training in connected speech processes (Experiments One to Five). In each experiment, it was hypothesised that if word-spotting training was beneficial, the group receiving that method of training would perform better, or no worse, on the post-test than the groups receiving the established training methods.

A pre-test – post-test design was chosen for this study in order to assess the efficacy of the word-spotting task in improving listening outcomes. While unlikely to be employed in the language classroom, such a design is widely used in training studies (as will be shown in Chapter 3) where a standard of comparison is essential.

---

1 Published in the author’s maiden name, Shanley.
1.3 Outline of the Chapters

Chapter 2 provides a background to the psycholinguistic processes underlying L1 speech segmentation by reviewing the literature on the multiple influences and cues involved. The use of language specific strategies in non-native listening is reviewed. Research showing some cues can be learned and others suppressed is then discussed in greater detail.

Chapter 3 provides a background to current approaches in the field of second language listening research. Previous perceptual (or ‘decoding’) training studies are evaluated and the benefits of the word-spotting training method are discussed.

In Chapter 4, stimuli creation for the word-spotting materials is discussed in depth, followed by a full description of the pilot experiment.

Chapters 5 – 9 present a series of five experiments, each with differences made to the research design, most notably the post-test used.

Finally, Chapter 10 summarises the results of the study and reviews the indications for the efficacy of the word-spotting task in training second language speech segmentation. Limitations of the study and suggestions for future research are also discussed.
Chapter 2

Speech Recognition and Segmentation

2.1 Speech Recognition

The recognition of words in continuous speech is made possible due to two mechanisms: 1) the parallel activation and subsequent competition of candidate words and 2) the use of acoustic cues to segment the speech stream. As we hear speech, we form hypotheses about what the component words in the stream might be. We match the sound sequences heard in the input to our stored lexical entries and consider multiple word candidates for each segment of speech, simultaneously (see McQueen, 2005).

Evidence for the activation of multiple candidates in parallel from the onset of the speech input comes from cross-modal priming experiments. For instance, in a Dutch study by Zwitserlood (1989), words semantically associated to both kapitaal and kapitein, (capital and captain, respectively) were primed by exposure to the fragment /kæpt/. If one should hear the sentence, “All the ideas in the universe can be described by words”, initially the words you and youth will be activated at the onset of the word universe, alongside all other words that commence with the sequence /ju/. These
words will fall away so that only more refined hypotheses remain as the word continues, such as *uniform*, *unicycle*, and *universe*.

Under connectionist models of spoken word recognition, word candidates are not only activated, but actively compete for selection. The competition between words is resolved through a process of inhibition of increasingly less likely candidates and enhanced activation for candidates receiving more support from the speech stream. The more strongly the favoured candidates are activated, the more the disfavoured candidates are inhibited. This process forms an integral part of the TRACE (McClelland & Ellman, 1986) and Shortlist (Norris, 1994) models.

At the point that the input satisfies the selection of *universe* in our example sentence, the word will be strongly activated, and the words considered previously that no longer match the input will be inhibited. However, words like *university*, *universal* and *universally* will still be possible candidates and will continue to receive activation\(^2\). The input will be analysed across the word boundary between *universe* and *can* before the other words receive strong enough inhibition to fall away.

Evidence for activation of candidate words across word boundaries comes from Gow and Gordon (1995), who showed that the meaning of *tulips* is primed when listeners hear the sequence *two lips* (the study being conducted with American English participants for whom the word *tulip* is pronounced without glide [y] insertion within the syllable /tu/).

It was once thought that candidates were activated in a strictly temporally linear, metaphorically 'left to right' fashion, as in the Cohort model of spoken word recognition (Marslen-Wilson & Welsh, 1978). However, there is evidence that the

\(^2\) The distinction between primary and secondary stress is ignored in lexical access by native English listeners (Cutler, 1986).
rhymes of words also stimulate activation of candidates, albeit less strongly than the onsets of words (Allopenna, Magnuson & Tanenhaus, 1998). The embedded words you and verse are also present in the input when we hear the words universe, university and universally and will be activated and compete with candidates activated from the onset of the word.

The competition between activated candidate words contributes to the segmentation of the string. As pairs of words increase in activation, others spanning the boundary between them will be inhibited (McQueen, 2005). Norris et al. (1997), have proposed that an additional mechanism, the Possible Word Constraint (PWC), penalises candidate words that do not coincide with possible word boundaries, and result, for instance, in a ‘word’ that consists simply of one or more consonants.

Higher-level, semantic cues do not have as strong an influence on this competition process as the phonetic detail of the speech stream. Norris, Cutler, McQueen and Butterfield (2006) showed that phonological forms are activated and take part in the competition process before conceptual representations are considered.

The notion that spoken word recognition relies on mechanisms of parallel activation of multiple lexical candidates is well accepted and is the basis for all speech recognition models to have been proposed over the last three decades (Cutler, 2012).

### 2.2 Segmentation Cues

But how are words selected for activation in the first place? The process of speech segmentation, whereby word boundaries in the continuous stream are posited, is an integral component of the procedures of word recognition and speech perception.

Multiple cues to word boundaries have been proposed and listeners have been shown to make use of many of them (Johnson & Jusczyk, 2001; Newman, Sawusch &
Wunnenberg, 2011; Mattys et al., 2005, Vroomen et al., 1988). Mattys et al., (2005) propose that there is a hierarchy of cue use, with listeners employing less-preferred cues (that is, those lower on the hierarchy) only if difficulty arises, for example when the speech signal is degraded. While the focus in this study is on the phonotactic and prosodic cues to segmentation, a brief overview of other proposed cues is warranted here.

2.2.1 Lexicality

It has been suggested that listeners use lexical knowledge to segment the speech stream. As previously discussed, awareness of what constitutes a possible word constrains segmentation. Listeners do not posit word boundaries that would result in remaining non-words (Hanulíková, et al., 2011; Norris et al., 1997).

In a cross-modal priming study, White, Melhorn and Mattys (2010) found that disyllabic possible word onsets, for instance corri, were stronger primes for full words, in this case corridor, when encountered in contexts where they followed real words, such as anythingcorri, than nonwords, as in imoshingcorri. This finding suggests that knowledge about the lexicality of sequences facilitates segmentation of the input.

The role of lexicality in speech segmentation is not an uncontroversial issue, however, and White et al.’s (2010) findings conflict with evidence that the sublexical process of segmentation precedes lexical access and selection of candidates for competition (Norris et al., 2006). In addition, in a word-spotting study that exploited potential neighbourhood effects (see Luce & Pisoni, 1998) for remaining nonsense syllables, Newman et al., (2011) found that the similarity of residual contexts to real words did not impact on the ease with which listeners detected target words.
2.2.2 Statistical cues

Prelinguistic infants are able to segment artificial streams of concatenated speech by learning the frequencies, and therefore probabilities, of different segment sequences. Sequences that occur within a word, such as /pri/ and /ti/ to form the word *pretty*, are statistically more frequent than those that occur across word boundaries, such as /ti/ and /bet/ in *pretty baby*. The term ‘sequential probabilities’ will be used here to refer to information within syllables, for instance /pri/ in the above example *pretty*, while ‘transitional probabilities’ will relate to information across syllables, such as between the /ti/ and /bet/ of *pretty baby*.

Saffran, Aslin, & Newport (1996a) conducted artificial language learning (ALL) experiments in which they presented 8 month old infants with two minute streams of four three syllable nonsense words, repeating in random order. There were no pauses between the ‘words’, nor stress cues. The exposure period was followed by a familiarisation-preference procedure. In the first experiment, infants had significantly longer listening times to unfamiliar nonwords, suggesting greater interest in them, than to the words presented in the ALL stream. In the second experiment the test phase contrasted ‘words’ within the artificial language and ‘part-words’ composed of the final syllable of one word and two syllables from the start of another word. Despite having received only two minutes of exposure to the transitional probabilities of the artificial language, the infants again displayed a listening preference for the more novel part-words. Both experiments showed that infants were able to find word boundaries using only the transitional probabilities between syllable pairs.

Statistical cues are also among the multiple cues used by adult listeners to segment continuous speech. Van der Lught (2001) used the word-spotting task, described in
Chapter 1, to study the use of sequential probabilities by Dutch adults in segmenting natural speech. The sequence *ga* is a common onset in Dutch, but the onset *geu* is rare. Listeners found words with common Dutch onsets (for example, *galg* ‘gallows’ in the string *hiengalg*) more easily than words with rare onsets (for instance, *geur* ‘aroma’ in *piengeur*). Dutch listeners, and potentially those of other language backgrounds, make use of sequential probabilities in finding word boundaries, at least in onset position.

### 2.2.3 Allophones

Systematic phonetic realisations of phonemes usually correspond to distinct word positions and can thus provide cues to word boundaries. English-learning infants as young as 2 months old are able to distinguish the allophonic variants of /t/ and /ɾ/ found in *nitrate* versus *night rate*, and hence have an implicit awareness of cues that may later assist in speech segmentation (Hohne & Jusczyk, 1994). At 10.5 months, but not 9 months of age, native English-learners are sensitive to the typical distribution of allophones within words (Jusczyk et al., 1999a). However, infants’ abilities to use allophonic cues in segmentation develop after their ability to make use of prosodic information and distributional cues.

Nevertheless, allophonic information plays an important role in adult speech segmentation. Newman et al., (2011) presented adult native English listeners with target-word-final word-spotting sequences in which the final consonant of the preceding nonsense context was either produced as a coda to that nonsense syllable (for instance, *veeф-apple* and *vuff-apple*) or as an onset to the vowel initial target word (as in *veeф-apple* and *vuhф-apple*). Irrespective of the length of the vowel, listeners had significantly faster reaction times and increased accuracy in finding the target word
when the preceding consonant was phonetically realised as a coda to the nonsense syllable and therefore matched the word boundary.

### 2.2.4 Prosodic cues

Native language prosody provides important cues for speech segmentation which are exploited early. Listeners use varying aspects of prosody as word boundary cues. Rhythmic segmentation strategies will first be discussed, followed by a strategy based on intonation.

#### 2.2.4.1 Rhythm

**Stress**

A stressed syllable is made with more respiratory effort than other syllables. In English, stress is most reliably perceived as increased duration, and often also as increased loudness and pitch (Ladefoged, 2001). Pitch is the perceptual correlate of the rate of vibration of the vocal folds, which results in a complex periodic waveform. The rate of repetition of this waveform is its fundamental frequency (F0; measured in Hertz (Hz)) (Johnson, 1997; Ladefoged, 2001, 2003).

The perceptual distinction between stressed and unstressed syllables (or strong and weak feet) in English is particularly clear, with unstressed vowels being reduced to one of a range of short, mid-central vowels. Reduced vowels are often referred to as ‘schwa’ and represented by the symbol /ə/ (Gussenhoven & Jacobs, 2005).

English features stress-based rhythm with an alternating strong-weak pattern (Abercrombie, 1967; Pike, 1945). English speakers use the prevalence of strong initial syllables (Cutler & Carter, 1987) to find word onsets. English-learning infants are able to apply implicit knowledge of the ubiquity of English strong-weak stress patterns to find word boundaries in speech input from the age of 7.5 months (Newsome &
Jusczyk, 1995) and continued experience with English speech input reinforces the efficiency of the strategy.

A seminal study by Cutler and Norris (1988) investigated the function of stressed syllables in the segmentation of English and included the first use of the word-spotting task. Listeners were presented with monosyllabic target words embedded in initial position in disyllabic nonsense strings. The target words contained short vowels and ended with a cluster of two consonants, for example, *mint* and the nonsense sequences started with a vowel, for instance, *-ayf* and *-esh*. They found slower response times for the recognition of target words in items containing two strong syllables, for example, *mint* in *mintayf*, than in items where the nonsense context contained a reduced vowel, for instance, *mint* in *mintesh*, where the vowel in the second syllable is schwa. The fact that target words in strong-strong nonsense strings were recognised more slowly than those in strong-weak strings suggests that the strong-strong words were segmented at the second syllable and the final consonant of the first syllable was erroneously assigned to it as an onset (resulting in *tayf*). It is important to note here that a separate experiment and consultation with trained phoneticians were conducted to ensure that the target words were not produced differently (that is, with positional allophonic variation) in the *mintayf* and *mintesh* conditions.

These results provide evidence that native English listeners segment continuous speech at stressed syllables. It is important to note here that, in lexical activation English listeners have been shown to make more use of contrasting vowel quality across strong and weak syllables, than the cues provided by the stressed syllable alone (Cooper, Cutler & Wales, 2002).
English stress cues are particularly important in first language acquisition and when the speech signal is degraded through faint speech (Cutler & Butterfield, 1992) or noise (Mattys et al., 2005; Smith, Cutler, Butterfield, & Nimmo-Smith, 1989) or contains inconclusive boundaries due to embeddings and novel words.

**Syllables**

French has syllable-based rhythm (Abercrombie, 1967), as does Korean (Stockmal, Moates & Bond, 2000; Yoon & Derwing, 2001). French listeners segment speech based on syllable rhythm (Mehler, Dommergues, Frauenfelder & Segui, 1981). They are faster to recognise targets that correspond with complete syllable shapes in the input than those that do not, irrespective of the size of the target. For instance, the target *pal* is found faster in the word *palmier*, where the /l/ forms the coda of the first syllable, than in the word *palace*, where the /l/ is syllabified as the onset of the second syllable. In contrast, the target *pa* is found in the word *palace* faster than in the word *palmier*. These results were not replicated when English listeners performed the same task with the same materials (Cutler, Mehler, Norris, & Segui, 1986), with English listeners showing no evidence of segmentation through syllabification. Importantly, when given English words, French listeners continued to segment the material based on syllable shape.

Kim, Davis and Cutler (2008) gave the same task (with Korean words) to Korean listeners and found that they also segment speech based on syllable structure. Listeners found the target *cam* faster in the word *campo*, where the target and first syllable correspond, than they did in the word *camay*, where they do not. Again, syllable structure did not influence the results, for example, the target *ca* was found more quickly in the word *camay*, than in *campo*. They then gave the Korean listeners the same French material as used in the Mehler et al. (1981) and Cutler et al.
(1986) studies and found that they segmented it in the same way as French listeners. The differing performance by English and Korean listeners on French material gives further evidence that segmentation strategies are based on the rhythmic structure of the native language of the listeners, not the language of the material presented.

**Morae**

The mora is the smallest unit of rhythmic structure. Japanese has mora-based rhythm (Kubozono, 1989; Port, Dalby & O’Dell, 1987), and Japanese listeners segment continuous speech based on the moraic structure of words (Otake, Hatano, Cutler & Mehler, 1993). In Japanese, the mora can be a single vocalic nucleus (V), a consonant + vowel (CV) or consonant + glide + vowel (CCV) sequence, or a syllabic nasal consonant in coda position (N). For instance, the Japanese name, *Honda*, has three morae: HO-N-DA (Cutler & Otake, 1994).

Using a fragment detection paradigm, Otake et al. (1993) found that the target *ta* was as easily detected in the word *tanshi*, where it corresponds to the initial mora, followed by a second mora /n/, as it was in the word *tanishi*, where it corresponds to the initial mora, followed by a second CV mora, /ni/. However, the miss rates for detection of another target, *tan*, in the word *tanishi* were very high. Here the /n/ which forms the final mora in the target *tan* does not correspond to the /n/ which forms the onset of the second mora in *tanishi*. These results confirm segmentation based on the mora, not the syllable, by native Japanese listeners. Otake et al. presented the same materials to English and French listeners and found that they did not present the same difficulty. Specifically, French listeners were shown to segment Japanese based on the syllable, as they do their native language. In contrast, Japanese listeners apply moraic segmentation to French and Spanish (Otake, Hatano & Yoneyama, 1996), for which syllabic segmentation is more suited. A subsequent
word-spotting study provided further evidence for moraic segmentation by Japanese listeners (McQueen, Otake & Cutler, 2001).

The Dravidian language Telugu has been shown to have similarities to Japanese and to also have moraic rhythm (Brown, 1991; Krishnamurti, 1998; Murty, Otake & Cutler, 2007). In a study similar to that of Otake et al. (1993), Murty et al. (2007) found that Telugu listeners’ response times in detecting target phonemes in moraic and nonmoraic Telugu contexts did not produce a consistent pattern. The authors state this may have been due to inadequate control in stimuli selection as a result of the lack of lexical databases on the Telugu language. Nevertheless, when Telugu speakers were presented with the Japanese materials used in Otake et al., they showed evidence of moraic segmentation similar to Japanese listeners. When Japanese listeners were presented with Telugu materials, they segmented them based on the words’ moraic structure.

The preceding discussion showed that native listeners segment speech according to the rhythmic structure of their first language. Rhythmic structures vary across languages. English has stress-based rhythm and a vocabulary with a predominance of initial stress. English speakers place word boundaries before stressed syllables. Korean and French have syllable-based rhythm and their native speakers use syllabic structure to segment speech. Japanese and Telugu feature mora-based rhythm and Japanese and Telugu listeners segment speech according to moraic structure. Korean and French are unrelated languages, with different syntactic structures and phonemic inventories. Telugu and Japanese are also unrelated. It is a shared rhythmic structure that leads these languages to be segmented in the same way.
2.2.4.2 Intonation

As noted earlier, French has syllable-based rhythm. French does not have lexical stress; that is to say, stress does not play a role in distinguishing words as it does in English, or even some other languages with syllable-based rhythm such as Spanish and Italian (Jun & Fougeron, 2002).

However, French does have sentence level salience differences, referred to as phrasal accent. Word-final syllables at the very end of phrases receive phrasal accent, perceptually realised as increased intensity and duration, and often rising pitch. To illustrate, in the sentence, *un gentil garçon chantait* (a nice boy was singing), the final syllable of *garçon* (boy) will receive increased emphasis as it is the last syllable of the Accentual Phrase (AP) (Jun & Fougeron, 2002).

As accented syllables occur at the right edge of the AP in word-final position, they provide reliable cues at which to locate word boundaries. Native French listeners have indeed been shown to exclude lexical competitors activated across word boundaries based on their knowledge that word-final boundaries equate to accented syllables in Accentual Phrases (Michelas & D’Imperio, 2010) and also Phonological Phrases (PP) (Christophe, Peperkamp, Pallier, Block & Mehler, 2004).

Seoul Korean – a regional accent of Korean, which as discussed earlier, is also a language with syllable-based rhythm – also employs pitch rise in the final syllable of APs (Jun, 1998; Kim, Broersma & Cho, 2012). According to a corpus study conducted by Kim (as cited in Jeon & Nolan, 2010), excluding those APs in final position in an Intonational Phrase (a higher phonological unit which may contain one or more AP (Jun, 2005)) 88% of APs feature a final rise.

Jeon and Nolan (2010) showed that Seoul Korean listeners use their knowledge of pitch rise in the final syllable of APs to segment continuous speech. Listeners were
exposed to pairs of ambiguous strings of digits. The strings were identical but for the location of the AP boundary, manipulation of which would result in a different number. Listeners heard one of four conditions: 1) original prosody maintained, 2) duration swapped between the utterances, 3) F0 (fundamental frequency – perceptually realised as pitch) swapped between the utterances, and 4) both duration and F0 swapped. They were asked to identify the second number (hence, the second AP) from the five-syllable string, and were presented with four alternatives to choose from on a screen. The alternative chosen revealed the way the listener had segmented the string. The effect of duration was found to be insignificant and only manipulation of the F0 resulted in a change in segmentation of the number string. The effect of the pitch rise cue was robust across different segmental contexts.

2.2.5 Phonotactic cues

Listeners also make use of phonotactic cues, the sequencing of speech sounds, to decide on word boundaries. For instance, Dutch listeners find the Dutch target word rond (round) faster in the nonsense sequence semrond than in sibrond as the sequence /mr/ is impermissible and must suggest a boundary, whereas /br/ is permissible (e.g. bruin, brown) and may be a word onset, making rond more difficult to segment from the string (McQueen, 1998).

The phonotactic constraints observed are often based on the Sonority Sequencing Principle (Jespersen, 1904; Selkirk, 1984), which captures the similarity of the segments constituting onsets and codas across many languages (Gussenhoven & Jacobs, 2005). This states that preceding the syllable nucleus, the acoustic intensity of consonant sequences should progressively increase and following the nucleus it should decrease. As sounds produced with an open vocal tract are most sonorous and
those produced with a total closure are least sonorous, vowels are at the top of the
sonority hierarchy and stops are at the bottom.

Listeners can also make use of sequences denoting more complex phonological
processes. Warner, Kim, Davis and Cutler (2005) found that Korean listeners were
sensitive to processes such as palatalisation, nasal assimilation and coda
neutralisation, and the resulting legality of certain coronal-/i/ sequences. Listeners
were asked to find /i/ initial words in word-spotting nonsense items where the
sequence across the word boundary was: 1) /詹姆/ – impermissible as a word boundary
cue as coda neutralisation would result in the voiced palatal affricate, /詹姆/, being
realised as an unreleased voiceless alveolar stop, [tʰ], in a word final context, 2) /ʤi/-
permissible as a word boundary, but more likely to appear syllable-internal, and 3)
/nni/ – permissible as a word boundary and always cues at least a syllable boundary.
As expected, the listeners found it most difficult to find /i/ initial words in the /詹姆/
boundary condition, followed by the /ʤi/ condition, with the /nni/ condition being
easiest of the three to segment. These findings indicate that Korean listeners are able
to use their knowledge of licit and illicit phonotactic sequences to locate word
boundaries despite the complexity of the underlying phonological alternations.

2.3 Second Language Speech Segmentation

Research on how listeners segment languages other than their native language reveals
that segmentation cues reside inside the listener’s mind, as part of their linguistic
competence, rather than within language structure. A large amount of research has
shown that listeners apply their first language segmentation strategies when listening
to non-L1 languages in which they have little or no fluency, much of which was

3 Transcribed in Warner et al. (2005) as /j/. The IPA transcription is used here.
discussed in the section on rhythmic segmentation strategies, above. The following sections outline further research on non-L1, and also bilingual\(^4\) and L2 segmentation, with a focus on prosodic and phonotactic strategies.

### 2.3.1 Non-L1 Listening

Additional evidence for language-specific speech segmentation strategies comes from artificial language learning (ALL) studies, in which cue use by listeners from diverse language backgrounds can be compared using the same artificial language input. The task of the listener is to segment a continuous stream of artificial words. The research demonstrates that listeners find it easier to segment novel languages when the word boundaries coincide with prosodic cues used in L1 speech segmentation.

Tyler & Cutler (2009) exposed English, French and Dutch listeners to a stream of artificial words containing either 1) pitch rise on the *initial* syllable of each artificial word, 2) pitch rise on the *final* syllable of each artificial word, or 3) no prosodic cue. They found that English listeners were assisted by a pitch movement cue on the first syllable of artificial words, while French listeners benefited from one on the final syllable. Cue use by these listeners in segmenting artificial speech conforms to the strategies they have been shown to use in segmenting their native languages. As discussed earlier, English listeners use knowledge of the prevalence of initial stress to find word boundaries, while French listeners make use of word-final accented syllables in Accentual Phrases. In keeping with their greater attention to suprasegmental cues (Cutler 2009), Dutch listeners made use of a pitch movement cue in both positions (but see Kim et al. (2012), as discussed below).

\(^4\) Here, bilinguals are defined as those exposed to two languages from birth or early childhood, who have native fluency in each.
In an additional experiment, a potentially language-universal cue of final lengthening was also observed, with English, Dutch and French listeners benefiting from a final lengthening, but not an initial lengthening cue; a finding also made earlier for English listeners by Saffran et al. (1996b).

Kim et al. (2012) also found L1-specific segmentation of an artificial language by Seoul Korean and Dutch listeners. The participants in this study heard a continuous stream of words with either 1) an initial pitch rise, 2) a final pitch rise, 3) final lengthening, 4) final lengthening and a final pitch rise, or 5) no prosodic cues. Korean listeners were able to find word boundaries more readily when final syllables carried a rise in pitch than when the words contained no prosodic cues. Recall that pitch rise is a consistent cue to the end of an Accentual Phrase (AP) in Korean (Jeon & Nolan, 2010; Jun, 1998). The Korean listeners did not benefit from a word-initial pitch rise, which only accompanies tense or aspirated initial stops in Korean and thus is not a consistent prosodic cue.

The final pitch rise cue did not assist Dutch listeners, for whom L1 phrase final intonation patterns vary. There was also no benefit received from the initial pitch rise cue, which Kim et al. (2012) suggested was due to the inconsistency of the trochaic (strong-weak) stress pattern in Dutch. However, Dutch listeners had responded to an initial pitch rise cue in Tyler and Cutler's (2009) study. Kim et al. suggest the difference in performance of the Dutch groups over the two studies may have been due to the difference in magnitude of the rise. In the Tyler and Cutler study, F0 was increased by 6 semitones, whereas in Kim et al. it was increased by 4.75 semitones. Kim et al. suggest the use of pitch movement cues by Dutch listeners in the previous study was a response at the acoustic-phonetic level, and not phonologically-determined. This analysis does not explain the lack of benefit gained from the word-
initial pitch rise cue by French listeners and from the word-final pitch rise cue by English listeners, who should surely have been just as receptive to an unmistakable phonetic cue; hence, the Dutch listeners’ performance in Tyler and Cutler’s study does indeed appear due to their greater awareness of prosodic cues.

Kim et al. (2012) found that both Korean and Dutch listeners responded strongly to the final lengthening cue, providing increased support for its universality. However, neither Dutch nor Korean listeners benefited from a cumulative effect of final lengthening accompanied by final rise in pitch. In a previous ALL study with Swiss French listeners, the combination of the two cues had also failed to enhance word boundary detection beyond the use of either cue alone (Bagou, Fougeron & Frauenfelder, 2002).

### 2.3.2 Bilingual Listening

Speech segmentation strategies are clearly constrained by the phonological structure of the listener’s native language. Questions naturally arise as to how bilinguals, particularly early bilinguals, segment their two languages. Do they have the perceptual flexibility to switch from one strategy to another in response to the language they hear?

Cutler et al. (1992) studied the segmentation strategies used by French/English bilinguals with native proficiency in both languages, as judged by native speakers. The participants had been exposed to the two languages from one year of age or earlier and used both on a daily basis. They were surveyed for the native language of each parent and their own preferred language. As none of the participants expressed a simple preference for either language, they were asked the question, “Suppose you developed a serious disease, and your life could only be saved by a brain operation which would unfortunately have the side effect of removing one of your languages.
Which language would you choose to keep?" The answer to that question was considered to be the participant’s dominant language.

The bilinguals were tested using exactly the same syllable monitoring procedures and materials employed in the monolingual studies by Mehler et al. (1981) and Cutler et al. (1986). The speed at which listeners could find CVC syllables, for instance bal and CV syllable sequences, ba, in words such as balance and balcony in the French material (and balance and balcony in the English material) was measured.

Recall that in the Mehler et al. (1981) and Cutler et al. (1986) studies, French monolinguals responded faster to French and English language stimuli, respectively, when the syllable structure of the target fragment was mirrored in the stimulus word, for example, ba in balance, but not in balcony or balcon. In contrast, the speed of responses by the English monolinguals (Cutler et al., 1986) was not affected by syllable structure.

The French-English bilingual results found in the Cutler et al., (1992) study did not correspond to previous findings for monolinguals on either the English or French language stimuli. The data was then organised into subgroups based on the listeners’ biographical information and a further analysis was conducted. Neither country of residence nor first language of either parent produced results akin to the monolingual results.

However, analyses based on the bilinguals’ dominant language produced an interesting finding. Those who were French-dominant produced results identical to those of French monolinguals on the French material, but those of English monolinguals on the English material. Those who were English-dominant produced the same pattern of results as monolingual English speakers on both the English and French language stimuli.
The results reveal that French-dominant bilinguals do not use their syllabic segmentation strategy on English words. They have learned that the French language is best segmented by syllables, but unlike French monolinguals, they have also learned that the use of syllabic segmentation is ineffective with the English language and do not apply it. Conversely, despite their extensive experience and native-level speaking proficiency with the French language, the English-dominant bilinguals do not use syllabic segmentation when they listen to French. In other words, French-dominant bilinguals suppress syllabic segmentation when listening to English, but English-dominant bilinguals do not apply syllabic segmentation when listening to French.

The question remained whether French-dominant French-English bilinguals would apply stress-based segmentation when listening to English, or whether only one rhythmic segmentation strategy was available to them, as appeared to be the case for English-dominant bilinguals. Cutler et al. (1992) tested a new group of French-English bilinguals. As in the previous experiment, the new participants were identified as being either French- or English-dominant by their choice of which language they would prefer to save from being lost. This second experiment used exactly the same materials and the word-spotting procedure of Cutler and Norris (1988); again the task of the listener was to detect target words amidst nonsense contexts. The test stimuli contained target words of either CVCC or CVC structure.

---

5 As Cutler et al., (1992) pointed out, the important question as to whether English-dominant French-English bilinguals would abandon their use of initial-stress cues to segmentation when listening to French also remained, but could not be answered in that study as French phonology does not permit a direct comparison to the Cutler and Norris (1988) study with English monolingual listeners. French phonology does not feature a stress system characterised by the juxtaposition of strong syllables with full vowels and weak syllables with reduced vowels.
joined to strong syllables (as in mintayve, thintayve) and weak syllables (mintesh, thintesh).

English-dominant bilinguals showed the same pattern of results as English monolinguals in the Cutler and Norris (1988) study. Target CVCC words preceding strong syllables (mintayve) were significantly harder to detect than those preceding weak syllables (mintesh), suggesting that the CVCC words in the strong-strong condition were segmented with the final C coda resyllabified as the onset to the nonsense context and then restored, while CVCC sequences remained intact in the strong-weak condition. There was no significant difference between the detection speeds of the stress shapes in the CVC condition (thintayve, thintesh) where no resyllabification was necessary.

On the other hand, French-dominant bilinguals did not show any difference in the ease with which they segmented the strong or weak contexts from the CVCC or CVC item sets. They did not apply stress-based segmentation to English despite it being the strategy best suited to the input. Neither group of bilinguals applied syllabic segmentation to the stimuli.

It seems that bilinguals suppress their dominant rhythmic segmentation strategy when listening to their non-dominant language, as they have had sufficient experience with the non-dominant language to have formed the awareness that their dominant-language strategy is ineffective and to avoid its application. Nevertheless, despite this extensive experience with and awareness of the non-dominant language, they do not acquire and apply the segmentation strategy best suited to it. The research of Cutler et al., (1992) suggests only one rhythmic segmentation strategy is available to the bilingual listener; even the highly proficient bilingual listener.
However, that strategy can be suppressed with sufficient experience of another language for which that strategy is not effective.

This finding was reinforced by a study on the perception of Spanish lexical stress by Spanish-French bilinguals, who were classified as either Spanish- or French-dominant depending on the primary language exposure in their earliest years (0-4 years of age). Dupoux, Peperkamp and Sebastián-Gallés (2010) found that in two short-term memory encoding tasks and a speeded lexical decision task, Spanish-dominant bilinguals performed in exactly the same way as native Spanish monolinguals, but the results attained by French-dominant bilinguals were merely comparable to those of French late-learners of Spanish, who had commenced L2 instruction at the age of 11. This study provides additional evidence that bilinguals do not perform at native-listener levels in both languages in aspects of prosodic processing.

### 2.3.3 Second Language Listening

It is clear that rhythmic segmentation strategies are language-specific. The fact that even highly experienced and adept bilinguals fail to segment their non-dominant languages in a way comparable to native listeners suggests that segmentation might pose a particularly serious challenge for the second language learner. Trouble may arise when a listener tries to apply their first language segmentation strategies to a second language with a different rhythmic and phonotactic structure. Research has shown that this is indeed what second language listeners do.

Japanese English-learning listeners have been found to apply their first language moraic segmentation strategy to English words (Cutler & Otake, 1994). In a phoneme monitoring study, Cutler and Otake measured the speed at which these listeners found two target phonemes in English words: the standard southern British English
vowel /ɒ/ and the alveolar nasal /n/. The English stimuli were matched to the
Japanese moraic structure contrast; for instance, /n/ appeared in candy in a moraic
context (in coda position) and in canopy in a non-moraic context (as an onset).

High miss rates for /ɒ/ and the speaker’s insertion of a glide between vowels in
sequence to avoid hiatus make the analysis of /ɒ/ problematic. However, Japanese
listeners were shown to both have lower miss rates for /n/ and to detect it faster in
moraic (candy) than non-moraic (canopy) contexts. This study provides further
evidence of the importance of the mora as a perceptual unit for Japanese listeners.
Unfortunately, it also suggests that the Japanese English-learners in this study were
applying an ineffective moraic segmentation strategy when listening to English.

Similarly, Tremblay and Spinelli (2014) found that, while native English learners of
French are able to make use of the distributional cues of their L2 in finding word
boundaries, they are also influenced by their L1 segmentation strategies. A visual-
world eye-tracking study was used to examine English and French listeners’ cue use in
locating word boundaries in liaison contexts in French. Liaison is the articulation of a
normally silent word-final consonant as a syllable onset preceding a vowel-initial
word, thus creating a misalignment of syllable and word boundaries.

Native English learners of French used duration as a cue to locate word boundaries in
liaison contexts, but native French listeners did not. In English, segmental duration
varies depending on word position (Ito & Strange, 2009; Turk & Shattuck-Hufnagel,
2000). In addition, as discussed previously, English listeners place word boundaries
before stressed syllables (Cutler & Norris, 1988), and one of the perceptual markers of
stress is increased duration (Gussenhoven & Jacobs, 2005; Ladefoged, 2001). The
misapplication of the English L1 cue in listening to French may prove advantageous in
some cases; however, Tremblay and Spinelli (2014) point out that duration is not a
consistently reliable cue to word boundaries in liaison contexts (accounting for French listeners not using it). As such, its use is problematic.6

The preceding discussion of second language segmentation focused on the difficulties faced by L2 listeners. However, a growing body of research demonstrates that L2 listeners can learn to suppress particular native language word boundary cues and to acquire those of their second language.

2.4 Learning Second Language Segmentation Strategies

2.4.1 Prosody

It seems some segmentation cues are more salient than others. As discussed, the research on the segmentation of artificial languages demonstrates that monolingual listeners are sensitive to pitch movement as a potent word boundary cue in unfamiliar languages. Research on second language segmentation reveals it is the lexical stress and accentual phrase boundary cues, and possibly the pitch movement associated with them, which lend themselves most to acquisition and suppression.

2.4.1.1 Suppression of Rhythmic Cues

We have seen that French-dominant French-English bilinguals are able to suppress their dominant-language syllabic segmentation strategy. Second language learners have also been shown to be able to suppress their native rhythmic strategies.

---

6 Tremblay and Spinelli (2014) propose two possible reasons for the use of duration as a cue to word initial boundaries in this study: the transfer of the duration cue to word boundaries from the English listeners’ L1, and the potentially universal response to durational evidence in non-L1 listening (as per Tyler and Cutler, (2009) and discussed previously). However, the evidence presented in Tyler and Cutler (2009), and also Kim et al., (2012) and Saffran et al., (1996b), was for the universality of a final lengthening cue, and no evidence was found in these studies for a universal initial lengthening cue. For this reason, the argument for a misapplication of the English listeners’ L1 strategy is more compelling than that of the employment of a universal cue.
In a word-spotting study, Hanulíková et al., (2011) found that highly proficient Slovak learners of German were able to suppress their L1 fixed stress segmentation strategy when listening to their L2. However, the learners were unable to suppress an L1 segmentation strategy based on implicit knowledge of viable residues when listening to German.

Likewise, using a word-monitoring paradigm, Tremblay, Coughlin, Bahler and Gaillard (2012) found that French-learning English listeners could inhibit the use of word-initial pitch rise as a cue to word boundaries when listening to French. Performance was linked to proficiency, with mid-level learners (mean of 8 years instruction) segmenting accented syllables as word-initial and high-level learners (mean of 12 years instruction) behaving like native French listeners in more often not segmenting accented syllables as word-initial. In this study, the learners were thought to use duration as the cue to French word boundaries and to be unable to use the French word-final pitch rise as a boundary cue. However, Coughlin and Tremblay (2012) believe this may have been due to the lack of sensitivity of the experimental task used.

**L1 and L2 Speech Segmentation by Spanish Listeners**

Bradley, Sánchez-Casas and García-Albea (1993) used the monitoring task designed by Mehler et al. (1981) (for use with French listeners) to study the way Castillian Spanish listeners use syllable structure in segmenting Spanish and English speech. Just as Mehler et al. had found with French listeners exposed to French input, Bradley et al. found that monolingual Spanish listeners showed a strong sensitivity to the syllabification cues of their L1; in this case, finding pal fast in palmera and pa fastest in paloma.
However, Bradley et al. (1993) found that monolingual Castillian Spanish listeners apparently abandoned that strategy and showed no sign of syllabic processing when listening to English. Surprisingly, in a subsequent experiment Bradley et al. also found that Castillian Spanish listeners who had lived in Australia for more than three years and were highly proficient in English did not show evidence of syllable-based segmentation of Spanish speech. Target syllables of both CV and CVC shapes were detected faster in CV (paloma) than CVC (palmera) stimuli. Bradley et al. make note that this pattern was also found by Cutler et al. (1986) for English listeners and English input.

A possible explanation for the seeming abandonment of Spanish L1 strategies might be that Spanish listeners recognise that syllable-based segmentation is ineffective when listening to English. This appears true of Spanish listeners with an abundance of experience with the English language, and also of monolingual Spanish listeners who, it must be noted, had been “coached... in the sound values to be encountered” in Australian English prior to the experiment (Bradley et al., 1993, p. 219).

2.4.1.2 Acquisition of Prosodic Cues

Just as Spanish listeners demonstrate sensitivity to the phonological structure of English in their suppression of L1 syllable-based strategies, they also show sensitivity to, and the ability to make use of, English stress cues.

In a phoneme monitoring study, Sanders, Neville and Woldorff (2002) found that Spanish English learners could learn to use English stress cues to segmentation. Near monolingual Spanish listeners did not use stress shape to locate target phonemes in sentences. However, Spanish English learners showed greater accuracy

---

7 Described as ‘bilingual’ in Bradley et al. (1993).
detecting targets in strong-weak than in weak-strong patterns, in correspondence to the frequency of the patterns’ occurrence in English (as per Cutler & Carter, 1987).

Spanish features syllable-based rhythm, but it also features lexical stress (Bradley et al., 1993; Soto-Faraco, Sebastián-Gallés & Cutler, 2001). Spanish words have a predominantly weak-strong pattern, with 87% of initial syllables being unstressed and only 13% being stressed (Sebastián-Gallés, Marti, Cuetos & Carreiras cited in Soto-Faraco et al., 2001). The English learners’ performance suggested an adaptation to the structure of their L2.

Sanders et al. (2002) queried whether the result was driven by those stimulus items in which the stress pattern was shared in English and Spanish. To ascertain whether this was indeed the case, the stimuli were divided for further analysis, into a set of shared stress patterns and another in which the stress pattern was specific to the English language. Sanders et al. found that the Spanish listeners displayed better performance with both sets of sentences, showing that they had learned and successfully exploited the most common English stress pattern in both cases.

It was noted earlier that a phoneme monitoring study showed that Japanese English learners apply moraic segmentation to English words (Cutler & Otake, 1994). However, Sanders et al. (2002) found that English-learning Japanese listeners also displayed greater accuracy detecting targets in strong-weak than in weak-strong patterns, suggesting they had learned the most optimal segmentation strategy for their L2.

The acquisition of prosodic cues has also been shown for French learners. In contrast to the Tremblay et al., (2012) study, Coughlin and Tremblay (2012) found that English speaking French learners can learn to use word-final pitch rise to find word boundaries in French. Participants were native English speaking, later-year university
students with seven years of instruction in the French language, and a control group of native French listeners. Coughlin and Tremblay used similar materials as had been used in Tremblay et al., but this time in a visual-world eye-tracking study examining participants' detection of monosyllabic noun targets in noun + adjective strings. The segmental boundary occurring across the two words could extend the first syllable to a disyllabic word and activate a competitor for the target noun. For instance, ce chat lépreux (this leprous cat) might activate the word chalet (cottage).

The noun + adjective sequence initiated one of two sentences. In one option the monosyllabic target was at the end of an Accental Phrase and carried a rise in pitch. In this version, the lexical competitor crossed an AP and was thus likely to be activated to a lesser degree than the target. In the other option, the target word and competitor were unaccented within an AP and could compete to a greater degree. Within each sentence type, there were two further options, manipulated through resynthesis to separate prosodic from durational effects. Participants heard only one version of each sentence. They also heard a similar number of distracter sentences.

The task of the listener was to click on one of four words presented orthographically on the screen before them - the target word, the competitor, or one of two distracters - as soon as they heard the word within each sentence. Their eye-movements were monitored.

In comparison to the native French listeners' performance, the cue use was delayed. The French learners' eye-tracking results also showed less prolonged fixation on the target compared to native French listeners, indicating a greater number of competitors were activated across word boundaries and Accental Phrases. Notwithstanding the limits to native-like efficiency, the results clearly showed that the native-English French learners were accurate in their detection of the target.
Importantly, the separation of pitch rise and durational cues in this study make it clear that they had learned to use the French word-final pitch rise as a cue to word boundaries.

**Implications**

Together, these studies, and those Artificial Language Learning studies described previously, demonstrate the potential for L2 learners to acquire prosodic cues to word boundaries. Further evidence comes from Suárez and Goh (2013), who showed that L2 lexical stress patterns are encoded in memory and used in word recognition. The adaptations made to acquire L2 word boundary cues based on lexical stress and accentual phrase boundaries may depend upon pitch movement, which is clearly a salient cue.

### 2.4.2 Phonotactics

Second language phonotactic sequences are notoriously hard for learners to master in production (Dziubalska-Kołaczyk & Zielinska, 2010) and perception (Dupoux, Kakehi, Hirosi, Pallier, & Mehler, 1999; Sebastián-Gallés, 2005). First language phonotactic constraints were shown to impact on the statistical learning of a novel artificial language (Finn & Hudson-Kam, 2008) and can impact on the success of second language segmentation, even for highly proficient learners. Nevertheless, while second language listeners may be unable to suppress existing phonotactic boundary cues, the perceptual system does appear to be flexible enough to learn new phonotactic segmentation mechanisms.

Weber and Cutler (2006) examined the use of phonotactic cues by English-learning German listeners in segmenting English language stimuli using the word-spotting task, with native English listeners as controls. The German participants were
students of English translation and interpretation, with an average age of L2 learning commencement of 11, and 15 years of subsequent instruction in English.

The stimulus material was constructed so as to fit into one of four boundary conditions, examples of which are presented here using the target word lecture. The phonotactic sequences either side of the ‘context-target’ word boundary suggested, 1) a boundary in English (tharshlecture), 2) a boundary in German (moycelecture), 3) a common boundary (moinlecture), or 4) no boundary in either language (gorklecture).

The target words in each condition started with either /l/ or /w/. The absence of the /w/ phoneme in German was not thought to be problematic to the design of the study as German listeners are known to map English /w/ to the /v/ of their German inventory.

While illegal, or impermissible, in English, the phonotactic sequences across the ‘English boundary’ condition (tharshlecture), did not suggest a word boundary in German. Many German words start with the /ʃl/ sequence, for instance, schlecht (bad), and schließen (to close), and the /ʃv/ sequence (in place of /sw/), as in schwanger (pregnant) and Schwester (sister). Likewise, the segments across the ‘German boundary’ condition (moycelecture) created illegal phonotactic sequences in German, but perfectly legal ones in English. /sl/ is a common word onset, for instance in slope and sly, as is /sw/ found in swing and swap. The ‘common boundary’ condition (moinlecture) contained sequences that would be impermissible syllable-internal in both languages, for example, /ml/ and /lw/, whereas the ‘no boundary’ condition (gorklecture), aligned phonotactic sequences permissible in onset position in both languages. To illustrate, the sequence /kl/ can be found in cluck in English and klingeln (to ring) in German.
The speed of detection and miss rates of the target words in the ‘no boundary’ (gorklecture) condition were compared with those of the other three conditions. As would be expected, the listeners’ performance was better in the ‘common boundary’ (moinlecture) condition. However, the English-learning German listeners also located targets more easily in the ‘German boundary’ condition (moycelecture) than the ‘no boundary’ condition, placing word boundaries between sequences such as /sl/, which, while not a permissible sequence in German, is a perfectly permissible onset in English. While the strategy did pay off in this case, following it when listening to the English language would be detrimental to efficient listening. This is particularly important given that these participants were training for careers as interpreters.

A more positive finding was that the German English-learners found it easier to segment the ‘English boundary’ condition (tharshlecture) than in the ‘no boundary’ condition (gorklecture). In this aspect, the L2 listeners demonstrated the listening behaviour of L1 listeners. The English-learning German listeners had developed enough expertise with the English language to acquire the English cue and impose a word boundary within the /ʃl/ sequence.

Al-jasser (2008) used the same methodology as that in the Weber and Cutler (2006) study to examine the learning of English phonotactic cues to speech segmentation by Qassimi Arabic English-learners. Al-jasser conducted explicit training in phonotactics, the details of which will be discussed in detail in Chapter 3, and the word-spotting task as the pre- and post-tests for this study. The Arabic participants were English language students at a Saudi university and had received nine years of English instruction. The group of Arabic listeners was placed into an experimental group, which would receive explicit training in impermissible English phonotactic sequences, and a control group.
Al-jasser lists the legal consonant clusters in the Qassimi dialect of Arabic as /dl/, /mr/, /bw/, /fl/, /kl/, /gr/, /kw/ and /θ/. English target words chosen started with /l/, /w/ or /r/. As in Weber and Cutler (2006), the word-spotting materials fit into one of four conditions, shown here with the target English word line as an example: 1) ‘English boundary’, where the segments over the context-target boundary contained sequences such as /dl/, which were impermissible in English, but legal in Arabic (veedline), 2) ‘Arabic Boundary’, where sequences would force a word boundary in QA, but not in English (veebline), 3) ‘common boundary’ (veetline), and 4) ‘no boundary’ (veefline). Experimental items were randomised between different lists. The lists also contained filler real words in final position that began with phonemes other than /l/, /w/, and /r/ and also some items that contained neither an Arabic nor English word. Like the Weber and Cutler study, performance in the ‘no boundary’ condition was compared to that in the other three conditions.

The pre-test results of the Arabic listeners were compared with those of American English listeners. As the test was conducted prior to training, the results for the Arabic control and experimental groups are combined and reported here as one group. Reaction times showed Arabic listeners found the English target words more slowly than native English listeners in every condition, and had significantly higher error rates in both the ‘English boundary’ (veedline) and ‘no boundary’ conditions (veefline) than the native English listeners.

Comparison among boundary conditions showed that the Arabic listeners made statistically fewer errors in the ‘English boundary’ (veedline), condition than in the ‘no boundary’ (veefline) condition. This suggests that the Arabic listeners had become sensitive to English boundary cues not occurring in Arabic over the course of their English education (and prior to the explicit phonotactic training given during this
study). However, detection times in the English boundary condition were not significantly faster than the ‘no boundary’ condition, indicating that there was some cognitive load involved in segmenting at the English boundary cue and automaticity had not been achieved.

While there was no significant difference in error rates for the native English and Arabic listeners in the ‘Arabic boundary’ condition *(veebline)*, the Arabic listeners were significantly more accurate in that condition than in the ‘no boundary’ condition *(veepline)*. This difference in accuracy suggests that Arabic listeners were unable to suppress native Arabic phonotactic cues to word boundaries when listening to English.

Administration of the pre-test and post-test was separated by nine weeks, with eight weeks of training in between. Post-test results showed no significant change in the error rates or reaction times of the Arabic control group. The Arabic experimental group’s performance in the post-test was still slower than the native English listeners’ pre-test results in all conditions. However, the Arabic experimental group showed significantly faster reaction times in the ‘English boundary’ *(veedline)* condition in the post-test than they had in the pre-test. In addition, there was now no significant difference in accuracy between the native listeners’ pre-test and the Arabic experimental group’s post-test performance in the English boundary condition. However, that finding would have been more compelling if a comparison with native listener post-tests had been possible. No change was shown in segmentation of the ‘Arabic boundary’ condition *(veebline)*.

Weber and Cutler (2006) and Al-jasser (2008) showed that German and Arabic learners of English, respectively, are able to successfully find word boundaries in L2 phonotactic sequences that would form allowable onsets in their L1, but were unable
to suppress the phonotactic boundary cues of their L1. However, it is important to note, that permissible English phonotactic sequences were not taught in Al-jasser’s study. L2 phonotactic cues may be highly responsive to training. The question remains as to whether training in permissible English phonotactic sequences may lead to the suppression of inapplicable L1 phonotactic cues.

2.5 Summary

The studies presented in this chapter provide evidence that the strategies employed in speech segmentation are predominately language-specific. The listener develops strategies best suited for the phonological structure of their first, or dominant, language. Listeners make use of the distributional probabilities, acoustic-phonetic cues, phonotactic restrictions, rhythmic structure, and intonational and phonological phrase boundaries of the language. There is also evidence that final lengthening may be a universal cue, with listeners of diverse L1s making use of it to find word boundaries in Artificial Language Learning tasks.

Second language listeners have been shown to apply the strategies appropriate for their L1 when listening to their L2, with negative consequences for the speed and accuracy of segmentation. However, several studies have documented the potential for listeners to learn segmentation strategies appropriate for their L2 and to suppress those of their L1. Pitch movement and phonotactic cues appear particularly amenable to acquisition. These findings raise intriguing questions about the potential for segmentation skills to be trained.
2.6 Can Second Language Segmentation be Trained?

The perceptual system has been shown to have the flexibility to respond to word boundary cues in a second language. New strategies have been acquired and, in the case of the Slovak German-learners described above (Hanulíková et al., 2011), old strategies suppressed through extensive exposure to the L2 over years of instruction. However, while cues have been learned, the delay in use of L2 phonotactic cues in the Al-jasser (2008) study and also the Accentual Phrase-final pitch rise in the Coughlin and Tremblay (2012) study suggest that L2 cue use may not easily be acquired to native-like processing efficiency. It follows that second language learners could benefit from focused training in the segmentation strategies best suited for the structure of the L2 in question.

Al-jasser (2008) has shown that phonotactic cue use may be improved through explicit instruction, which is a welcome finding, but segmentation training methodology should have further goals. It is important to note that the pre-test results of the study revealed that the Arabic participants had already developed the ability to segment at English boundaries. The study simply improved upon that knowledge. It would be more beneficial to develop a training program that led to the acquisition of new word boundary cues and also the suppression of L1 strategies. It is important that a segmentation training program also result in greater automaticity. Furthermore, Al-jasser used explicit training, but as will be discussed in the following chapter, there is some evidence that implicit training may lead to more native-like language acquisition.
Chapter 3

Training L2 Speech Segmentation

3.1 Introduction

The purpose of this chapter is to review the literature on second language listening teaching that has relevance to speech segmentation. It begins by discussing traditional classroom approaches to listening teaching and goes on to describe approaches that incorporate psycholinguistic research. Two differing approaches to the teaching of speech segmentation will then be discussed: decoding and metacognitive approaches. Unfortunately, there is little published research on attempts to train second language speech segmentation. The available research will be discussed, alongside training studies on the teaching of other relevant decoding processes (traditionally referred to as 'bottom up' processes).

3.2 Background

Traditional classroom approaches to second language listening have focused on presenting listening passages and asking questions to ascertain a student's level of comprehension. The comprehension approach typically involves three stages: pre-listening, while-listening and post-listening (Cauldwell, 2013; Field, 2008b; Richards, 2005). In the pre-listening stage new vocabulary is often explained. Students are
given a background context to the scenario and are sometimes asked to predict what they may hear in the listening passage. The while-listening stage is often composed of an extensive listening phase, during which students listen for general content, including who the speakers are and what they are speaking about; and an intensive listening phase, in which students listen for particular details in order to answer specific questions (Field, 2008b). Finally, at the post-listening stage, teachers often use the information or grammatical structures featured in the listening passage to form the basis of an extension activity. This may be a role play, or perhaps a discussion on the meaningful information derived from the listening passage. Sometimes the focus of this stage will be the grammatical features used in the passage. If vocabulary was not taught in the pre-listening stage, it is usually also taught at this point. Such an approach is generally considered beneficial in consolidating new language (Cauldwell, 2013), which may include vocabulary learning or mastery of syntactic and pragmatic structures.

Some writers have challenged the comprehension approach. Richards (2005) has questioned how a focus on comprehension can facilitate learning. Both Field (2008b) and Cauldwell (2013) have criticised the approach as testing listening, rather than teaching it. To elaborate, the task of the listener is to understand the passage and answer questions about it, but listeners are not provided with instruction or training to improve their comprehension. They also point out that the speech signal is not given importance beyond its role in carrying meaning. Likewise, the need for listeners to acquire processes in order to decode the signal and thereby match it to words (Field, 2008b) is not prioritised (and possibly not understood). Listeners are expected to acquire decoding skills through “osmosis” (Cauldwell, 2013, p. 252); that is, through extensive exposure to listening materials that are scanned for meaningful content. Clearly, any exposure to word boundary cues under this approach is
incidental, and speech segmentation skills are likely to develop slowly in response to this exposure if at all.

An early alternative approach centred on the skills required for successful L2 listening. The skills training approach divided the behaviour of listening into smaller contributory actions and provided practice at performing each of these in isolation before gradually recombining them (Field, 2008b).

Richards (1983) proposed a taxonomy of necessary micro-skills, some of which were directly relevant to speech segmentation: the abilities to discriminate sounds and to recognise reduced forms, stress patterns and the rhythmic structure of English. He also listed the ability to distinguish word boundaries, but he placed more emphasis on the role of grammatical and contextual knowledge in the segmentation of discourse than on that of phonetic and phonological competence in the segmentation of speech.

He considered segmentation of syntactic structure most important, as it enabled listeners to group speech into constituents and from there to identify the propositional, or meaningful content of the message. Grammatical competence was seen as essential to listening competence. However, in his view, contextual knowledge could bypass the process of constituent identification.

While Richards was influenced by psycholinguistic research by Clark and Clark (1977), Field (2008b) points out that his approach was not truly psycholinguistic in nature. The micro-skills he proposed were intuitive, and not based on empirical evidence. They also refer to a conception of abilities required in listening, rather than a set of processes employed.
There is a growing body of second language teaching literature that acknowledges psycholinguistic research into the processes used by native listeners, and the phonetic details of the speech stream in L2 listening. However, not all writers provide suggestions on how to teach these processes or details. For instance, Rost (2011) discusses knowledge gained from psycholinguistic research on how language learners listen to their L2 and provides much description of decoding processes. He also makes recommendations for further psycholinguistic research. In spite of that, he does not incorporate such research into recommendations for teaching, beyond arguing that listening materials should feature authentic input, the main characteristic of which, according to Rost, is ‘genuineness’. It is colloquial, spontaneous and preserves natural connected speech processes. However, Rost also discusses the virtues of ‘simplification’, including “overtly marking word and phrase boundaries by slowing down or exaggerating speech patterns” (p. 172). Such an approach is unlikely to provide listeners with the practice required to segment authentic, continuous speech in real-time.

However, there are other approaches emerging within the literature that are psycholinguistic in nature and do also encourage teachers to provide instruction and practice in decoding processes. While it must be acknowledged that there are differences in the foci and goals of the researchers involved, their work does share important features and will be discussed here as decoding approaches.

3.3 Decoding Approaches

In this thesis, the term decoding approaches refers to L2 listening pedagogy that focuses on the accurate matching of acoustic-phonetic input to words. Examples can be found in Hagen (2000), Brown and Kondo-Brown (2006), Field (2008b), and Cauldwell (2013). Like the skills training approach, decoding approaches view
listening as a behaviour that can be divided into smaller processes for extensive practice in isolation. However, in contrast to the skills training approach, the identification of such elements is informed by psycholinguistic research (Field, 2008b). This research also inspires goals, targets and methods.

The core goals of decoding approaches are generally to improve the accuracy of lexical access and to build automaticity (Brown & Kondo-Brown, 2006; Cauldwell, 2013; Field, 2008b). An important aim is to provide students with exposure to and practice with the continuous stream of authentic speech. Listeners are trained in the skills used to matching the fragmentary contents of the speech stream to words and to find word boundaries. They are also trained to make use of prosodic information in both lexical access and the identification of syntactic structures and pragmatic intentions.

Word recognition is an important target. Activities are suggested that raise awareness of the difference between the citation form of words (that spoken in isolation) and the way those words are realised in authentic, continuous speech. As will be discussed further, in some cases processes are explicitly taught, including reduction, assimilation, and intrusion, but in others the focus is firmly on the output of these processes. Other activities are designed to provide focused exposure to the forms that result from these processes and controlled practice in recognising such forms in short strings of speech.

The continuous nature of speech, itself, is also a focus. For instance, the linking of words can cause intrusive sounds to appear between them. Native English speakers often resolve hiatus (the occurrence of two vowels in sequence, as in when a vowel-final word is followed by a vowel-initial word) by producing an intrusive [r] between the vowels. To illustrate, Canada aided may become Canada raided (Tuinman, Mitterer & Cutler, 2012). Within the decoding approach to L2 listening instruction,
activities are designed to raise listeners’ awareness to this phenomenon and to provide controlled practice in achieving an L1-like perception of the string (see for instance, Brown & Kondo Brown, 2006).

The training goal in such cases is not to have the listener perceive the sounds ‘correctly’. The problem for the L2 listener is that they already do accurately perceive the contents of the speech stream. The goal is for the listener to be able to ‘ignore’ or suspend that accurate perception in order to arrive at an interpretation akin to that of the native listener, who is oblivious to such processes.

Decoding approaches seek to achieve such improvements in L2 listening through guiding the learner through controlled to automatic processing (Brown & Kondo Brown, 2006). Published listening teaching and learning activities generally take the form of short exercises targeting a specific feature of the speech stream. Field (2008b) refers to these as “small-scale classroom listening exercises”.

Decoding approaches make use of pre-listening, while-listening and post-listening stages in listening sessions, with a focus on the development of decoding processes rather than the completion of comprehension exercises. Although the focus within decoding approaches is on teaching listeners the processes or activities that will help them match the speech signal to words, the importance of meaning is also acknowledged. Decoding approaches do not recommend a focus on the phonetic and lexical levels of listening to the exclusion of meaning-building instruction and activities. Both are encouraged as part of an integrated approach.

A more detailed discussion of the ideas and teaching programs published in Cauldwell (2013) and Field (2008b) will be discussed in the forthcoming sections.
3.3.1 Cauldwell’s Spontaneous Speech Model

Cauldwell (2013) calls for a focus on teaching the phonetic form of spontaneous, unscripted, authentic English: the difference between the native (or expert) listener’s perception of what has been said, and the information that is actually present in the speech stream. The focus here might be on connected speech processes; however, Cauldwell’s is not a process-oriented approach, but one centred on the output of those processes.

Attention is drawn to the authentic ‘soundshapes’ of words; for instance, the reduced form of function words. Much of this reduction takes place in what Cauldwell (2013) calls the ‘squeeze zone’, which he explains as being strings of both stressed and unstressed syllables between the intonational peaks in a speech unit. He emphasises that syllables within the squeeze zone can sound dramatically different to the citation forms of the words. The perception of words within this zone is a major focus of Cauldwell’s suggestions for teaching.

Cauldwell’s (2013) practical suggestions follow Buck (cited in Cauldwell, 2013). These are: a) raise awareness to the problem, b) discuss it explicitly, c) explain how pronunciation changes in connected speech, and d) give examples. Much importance is placed on using unscripted materials in these activities, in short extracts of two to ten seconds’ duration.

The true contents of the speech stream are taught through pronunciation activities, metaphor, controlled ‘noticing’, and discussion. Cauldwell (2013) recommends teachers start with pronunciation. He suggests students be encouraged to manipulate prominence in sentences, and generate their own suggestions for reduced vowels, then reduced consonants, in their own speech. Cauldwell encourages the use of metaphors to explain complex concepts – descriptions of the physical, tangible and
sensory to explain the abstract - instead of explicit rules and tendencies. This is similar to the way speech pathologists and pronunciation teachers (for instance, Kimble-Fry, 2001) explain difficult phonological concepts and required oral behaviours. Indeed, Cauldwell is also an active pronunciation teacher and researcher and cites an influence of his former mentor, David Brazil, a pioneer in the teaching of pronunciation and intonation (see Cauldwell, 2012).

Suggested listening exercises include playing groups of frequent forms, such as “where there were”, and “of what I was going to do” in order to show how they are realised in repeated forms in a spontaneous text. It is also recommended that content words that appear more than once be played to highlight their different ‘soundshapes’. Cauldwell (2013) also suggests that after hearing a longer listening passage, students be given the transcript and asked to follow it as they listen again. Students should then be asked to mark difficult areas for classroom review.

In addition to classroom exercises, Cauldwell (2013) recommends making use of what he calls “hi-tech approaches”. These include CD-ROMs and apps for smartphones and tablets. Cauldwell provides suggestions as to current (at time of publication) examples, and his own CD-ROM and soundfiles are available on his website, www.speechinaction.org.

One criticism of Cauldwell’s (2013) approach is that he claims “the spontaneous speech model cannot be taught by rule and example” as “none of its features are under control of the listener, the learner, the teacher or the textbook writer” (p. 20).

It is true that words and formulaic phrases can be said in a multitude of different ways and, as will be discussed further in this chapter, implicit training does have great value. However, a consequence of Cauldwell’s statement is concerning: if listeners are not expected to be able to even generalise across examples, there is little point in
repeatedly exposing learners to different instances of the same set of words, as Cauldwell himself suggests in his list of proposed activities.

As shown above, Cauldwell (2013) is centred on the surface output of connected speech processes. As such, beyond exposure to assimilations and reductions across words, his approach does not explicitly cover speech segmentation. The problem of L2 speech segmentation has been recognised only in recent years. Consequently, the value of L2 speech segmentation training has received scant attention in the field of second language teaching and learning, and even less has been written on pedagogical suggestions for teaching it. One exception within the decoding approach is the work of John Field.

3.3.2 Field’s Process Approach

While the skills training approach divided listening into a set of sub-skills, and Cauldwell’s (2013) approach divides it into a set of outputs for listeners to become accustomed to, Field (2008b) divides listening into a set of contributory processes drawn from psycholinguistic research. Field has had particular influence in bringing psycholinguistic literature on speech perception to the awareness of the field of second language acquisition (for example, Field 2003; 2005; 2008b).

Field’s (2008b) goal is for the second language listener to build automaticity in employing the mechanisms involved in successful L2 listening, stating that the process approach is “a movement towards associating the goals of L2 skills training more closely with the processes underlying L1 performance”. Field’s list of decoding processes include those at the segmental, syllable and word levels (for instance, identifying consonants and vowels, matching weak syllables and function words, and identifying words which are not in their standard forms), and also those at the syntactic
and intonational levels (including, *anticipating syntactic patterns*, and *using intonation to support syntax*, respectively).

As an alternative to merely testing listening through comprehension exercises, Field (2008b) suggests listening teachers take a diagnostic approach in becoming aware of the reasons for failure in understanding and offering a set of remedial exercises to address problems. He advocates controlled exposure to features of the L2 input that pose difficulties for learners, repeated practice in using the processes employed by L1 listeners, and further instruction and practice in strategies to compensate for remaining decoding deficiencies. According to Field, in order to develop automaticity, the difficulty of listening materials should be varied along the gradients of speech rate, number of speakers, accents, and speaker fluency (fluent and planned versus dysfluent and spontaneous).

### 3.3.2.1 The Process Approach to Teaching Speech Segmentation

In contrast to most in the field of L2 listening pedagogy, Field devotes considerable attention to speech segmentation. For instance, his list of decoding processes used in listening includes ‘*working out where words begin and end in connected speech*’ (Field, 2008b).

Field (2008b) provides many practical recommendations on the training of English speech segmentation processes. For example, teachers are suggested to raise awareness of the potential for incorrect segmentation outcomes by asking students to dictate ambiguous sentences. Having the sentences end in unexpected ways encourages students to revise their segmentation hypotheses. For instance “the way to cut it” is likely to be transcribed as “the waiter cut it” due to a lack of awareness of the reduction of vowels in function words to schwa. However, the dictation of the
sentence ending, “is like this”, may prompt the student to rethink their original segmentation of the string.

Teachers are also advised to raise awareness of the weak syllables in prefixes and suffixes by providing students practice in recognising these in short extracts of three or four sentences. When the extracts are played a second time, students are asked to transcribe the words containing prefixes or suffixes. Field (2008b) also recommends a form of dictation (which he calls ‘standard segmentation practice’), in which students build up a transcription of authentic speech over multiple repetitions and then compare their answers to other students.

Other approaches serve as introductions to an explanation of the L2 segmentation system. For instance, Field (2008b) suggests teachers dictate an unbroken string of non-words that neatly permit use of the dominant strategies used in segmentation by native English listeners (which, as discussed in Chapter 2, are initial stress and illicit phonotactic sequences across word boundaries). After reviewing the students’ responses the teacher would explain the cues used by L1 listeners in locating word boundaries. A related approach sees the teacher dictating a short series of sentences that contain words above the students’ proficiency level. The teacher is advised to say, “You won’t understand everything, but guess how many words there are”. This serves as a springboard to an explanation of the segmentation system and a subsequent examination of the sentences in further detail. In both approaches, attention is drawn to the differences between the L2 listeners’ segmentations and those most likely to be made by native-English listeners.

Teachers are also provided with other suggestions on how to demonstrate and give listeners practice in using the cues native English listeners have been shown to use in finding word boundaries. For instance, a training activity might focus on word stress.
Field (2008b) suggests the teacher read short sentences aloud, while students follow along and mark stressed syllables on a transcript. The teacher would then give explicit instruction on the role of stress in English and its dominant location. This is a similar approach to that taken by Al-jasser (2008) in teaching Arabic listeners to make use of English phonotactics to find word boundaries.

Field (2008b) stresses the importance of advising listeners to treat the outcomes of attempts to use L1 cues as provisional only, and to be prepared to make adjustments as needed. This recommendation follows Field’s (2008c) finding that L2 listeners are less likely to revise original segmentation hypotheses than native English listeners.

Another pedagogical approach to the development of L2 speech segmentation skills exists in the literature. However its aims are quite distinct from John Field’s process approach and those of decoding approaches generally.

### 3.4 The Metacognitive Approach

The *metacognitive approach* is centred on the training of sequential, reflective listening strategies with a contextual, or ‘top-down’, focus (Berne, 2004; Chamot, 2005; Vandergrift, 2004). It is an elaboration on the traditional comprehension approach. Planning, monitoring and evaluation activities are incorporated into listening lessons. The planning stage includes previewing task demands and setting comprehension goals. During the monitoring stage, listeners review their understanding; for instance by comparing current and prior sources of knowledge. Finally, listeners evaluate the success of their listening by, for example, rating their level of comprehension, and judging the appropriateness of the strategies used (Vandergrift & Goh, 2012).
3.4.1 The Metacognitive Approach to Teaching Speech Segmentation

Vandergrift and Goh’s research (2012; see also Goh, 2000) reflects influences of Anderson’s (1995) proposal that listening comprehension has three interdependent phases: 1) perceptual processing, in which decoding of the acoustic-phonetic properties of the speech stream occurs, 2) parsing, in which potential word candidates are activated and integrated into chunks of propositional content, and 3) utilization, when the meaningful units are interpreted and monitored for congruency using contextual knowledge. This is not a sequence of disparate phases; the phases are integrated by bidirectional information flow between decoding and contextual processes. The mechanisms involved in speech segmentation commence in the perception phase and are completed in the parsing phase. Contextual factors are thought to exert the greatest influence on the speed and precision of speech segmentation and lexical activation.

Vandergrift and Goh (2012) review several speech segmentation training strategies, including decoding skills-based dictation and cloze (or ‘gap filling’) exercises. However, their focus is firmly upon the development of metacognitive strategies around planning, monitoring and evaluation of listening attempts; for instance, pre-listening planning activities that trigger contextual knowledge and allow prediction.

Vandergrift and Goh (2012, p. 147) assert that semantics has primacy in speech segmentation, claiming it to be “the most salient cue in perception”. (A similar claim is made by Mattys et al. (2005)). Yet, as discussed in Chapter 2, pitch movement is a highly salient and robust cue which has been shown to be readily adapted to analogous cues in the listener’s L2, and L2 phonotactic cues also appear to be readily learned.
3.5 Top-Down or Bottom-Up?

Within the second language listening literature, including that by Vandergrift and Goh (for instance, 2012; and see also Goh, 2000) there is a concern with striking the right balance between what is traditionally referred to as 'top-down' and 'bottom-up' processes, (but here as contextual and decoding (or perceptual) processes, respectively). Researchers often fail to acknowledge that conceptual and perceptual mechanisms are interdependent (Field, 2008a), and champion the former to the neglect, or the condemnation, of the latter. Some researchers have a distrust of listening pedagogy with a purely decoding focus, due to the concern that an emphasis on these processes may lead to a listener’s over-reliance on ‘word-by-word’ strategies and a consequent failure to acquire higher-level processes. Such viewpoints, however, are inconsistent with the evidence summarised in Chapter 2, which indicates instead that accurate speech perception is vital for higher-level processes to function successfully.

Some fear that a focus on the training of decoding processes will encourage lower-proficiency listeners to attempt to understand every word in the input, a strategy believed to interfere with the reception of contextual information that could fill gaps in comprehension. Vandergrift and Goh (2012) state that remediation of listeners’ segmentation difficulties must employ meaningful connected speech, so that the context of the text can assist listeners to identify otherwise unintelligible sounds. However, there is evidence from psycholinguistic studies (see Cutler, 2012; C. 12) that speech processing does not require feedback from semantic context influencing pre-lexical processing. Of course, context may provide the impetus for a re-evaluation of erroneous selection of lexical candidates, but this should occur only after candidate
selection, not during perceptual processing. In contrast, the improvement of pre-lexical mechanisms through instruction in decoding processes may have a substantial impact on real-time processing.

Further, an abundance of research shows that low-proficiency listeners actually make more use of contextual ‘top down’ information than high-proficiency listeners (see Field (2004) for a review). Field played low-proficiency English-learning listeners sentences where the final word was unexpected given the semantic context of the sentence and also infrequent and unlikely to be known at the listeners’ proficiency level (for example, “They’re lazy in that office; they like to shirk”). The listeners’ task was to write the final word of each sentence heard. Field found that the listeners did not attend to the segmental content of the words and, instead, substituted known words – for example, in the above sentence, work. Listeners’ responses were both phonologically inaccurate and semantically inappropriate. Still, the association between lexical items within the sentences and those responses (for instance, office and work), revealed that contextual information was relied upon, here actually to the detriment of perceptual processing.

Furthermore, Tsui and Fullilove (1998) found that low-proficiency listeners relied on higher-level, contextual processes as a compensatory strategy due to poor perception skills. They examined the results of the listening component of public English examinations in Hong Kong over a period of seven years to ascertain whether there were a link between ‘top-down’ and ‘bottom up’ processing, and listening proficiency. Tsui and Fullilove found that when the schema, or cognitive framework, around a test question was incongruent with the following linguistic input - for instance, the question asks, “What saved the estate from burning down?” - and the subsequent interview with a fireman mentions how lucky it was that the wind changed direction -
the lower-proficiency listeners did not have the decoding skills to find the correct answer. Instead, they relied on contextual processing.

Tsui and Fullilove (1998) stressed the importance of decoding processes in listening comprehension, concluding that low-proficiency learners need to become less reliant on contextual processes and more skilful at rapid and accurate perception of the input. The evidence is in stark contrast with Goh’s (2000, p. 70) erroneous assertion that, “Learners who repeatedly complain about low-level perception problems probably do not engage sufficiently in top-down processing”.

### 3.6 The Training of Decoding Processes

Field (1998) points out that there is little evidence to suggest that ‘top down’ metacognitive training is useful, with many training studies producing ambiguous results. In contrast, several studies employing the training of decoding processes, including speech segmentation and word recognition, have shown the decoding approach to be effective.

#### 3.6.1 Dictation

Dictation has long been suggested as a training method in listening instruction (for instance, Field, (1998, 2003); Morley, (1977); Nation and Newton, (2009); and Ur, (1984)). While it must be noted that Jafarpur and Yamini (1993) did not find evidence for an improvement in language skills, including listening comprehension, after 60 dictations given over a semester, dictation has more recently been empirically shown to be an effective method of improving listening skills.

Kiany and Shiramiry (2002) compared the listening comprehension scores of an experimental group who had been exposed to eleven sessions of dictation in addition to the listening exercises from their English text book, to a control group, who had
only received the text book listening lessons. The study was conducted with 60 male elementary English-learners in Iran over one term (20 sessions).

The English text book used in lessons for both groups was *Headway Elementary* (Soars & Soars, 1993). In the dictation sessions, listeners were first told the topic of the aural text. They then listened to the text in its entirety without any pauses. The tape was then played again and paused after each meaningful chunk so that listeners could transcribe what they heard. Subsequently, the tape was played in its entirety again to allow listeners the chance to check their responses. Finally, the listeners checked their responses against a written transcript while listening to the text once more.

A comparison between pre- and post-test results on a battery of elementary listening tests showed significantly better listening comprehension in the group that had received the extra dictation training than in the group that had received only the textbook lessons.

Note that the participants in this study were assigned to experimental and control groups based on existing class groups (two classes in each training group), not randomly. However, all four classes were taught by the same instructor (one of the researchers) and t-tests conducted on the pre-tests results showed no inter-group differences in language proficiency or listening comprehension.

### 3.6.1.1 Computer Assisted Dictation Training

Despite the interest in computer assisted language learning (CALL) generally, there is little research on its value for training listening skills. Hulstijn (2003) described multimedia software designed to promote the automaticity of L2 listening mechanisms. Listeners can use the 123LISTEN program to play full aural texts, fragments of texts with delayed subtitles, or fragments with simultaneous subtitles. Hulstijn predicted that the most value for L2 listeners would be gained from playing
the fragments as many times as required to be able to reconstruct the utterances before seeking feedback from the subtitles, but the efficacy of the program has not been evaluated empirically.

Matthews and O’Toole (2013) recently conducted an implicit training study utilising a web-based program in an attempt to train word recognition. The program bears a certain resemblance to the 123LISTEN software (Hulstijn, 2003), but contains more features and is essentially a listener-directed dictation task. In the study, English language learners enrolled in a South East Asian university received approximately 40 minutes of training per session for five sessions. Training materials consisted of five monologues with mean duration of 32 seconds and mean number of 66 words each.

In each session a full monologue was first presented aurally. Participants were then able to listen to sections of the monologue up to ten times and attempt to reconstruct the section by typing in an on-screen field. Each attempt was recorded by the software. Once they had made their final attempt, participants received onscreen feedback in the form of a comparison between their attempt and the target utterance. Errors and omissions were highlighted in red. Participants could then replay the monologue and view the text simultaneously.

The pre- and post-tests consisted of three different measures: 1) partial dictation tests of 20 sentences with missing target words, 2) paused dictation tests, in which a 30 sentences were played and paused for participants to write down the last word of the sentence, 3) a monologue dictation, which followed the same format as the training, but for the fact that participants were allowed just three opportunities to listen to each section and transcribe it.

Statistically significant improvements were made on all three measures. However, the study had important limitations. First, there was no control group and, as the authors
themselves note, improvements may have been due simply to learner maturation. Second, participants were given an alphabetised list of words used in the listening program, organised into grammatical or semantic categories, before commencing the training. This made the training activity one of word recall rather than recognition. The preview of words used in the training may be useful in earlier stages of a program, but listeners probably need to be provided with more authentic listening challenges if improvements are to generalise to real speech. Finally, and most importantly, the same test sentences and target words were used in the partial and paused components of the pre- and post-tests, with only the order of item appearance differing between the tests. As there was no control group in this study, the repetition of test items in the partial and paused dictations makes it hard to interpret the results of those post-tests. Improvement may simply be due to the additional opportunity to hear the sequence. Different texts were used in the full dictations, but of the 37 content words in the pre-test and the 39 in the post-test, five words were repeated in identical form and an additional four words shared the same root with derivational or inflectional suffix changes.

3.6.2 Explicit Training in Connected Speech Processes

Other researchers have sought to improve second language listening skills through the teaching of connected speech processes. The citation form of words differs markedly from their pronunciation in connected speech (Ingram, 1989), yet these processes are rarely taught in the second language classroom. Brown and Hilferty (1986) taught Chinese graduate students in Guangzhou the reduced forms of a selection of English words, with a focus on American English. The forms were taught as individual words, but reinforced through dictation practice throughout the training session. The lessons were given for ten minutes a day over four weeks. Three different post-tests were used to measure success: The Bowen (1976) Integrative
Grammar Test (which requires listeners to transcribe reduced forms), a multiple choice listening comprehension test and dictations. Significant differences were found between the means of the treatment and control group (which had received only minimal pair drills) in the Integrative Grammar Test and dictation. There was no significant difference observed across scores in the listening comprehension test.

In a more recent study, explicit training in the connected speech processes of French resulted in a reduction of inaccurately transcribed words in dictation activities (Kennedy and Blanchet, 2013). Training was provided to a linguistically diverse group of French second language learners in Quebec over 15 weeks. The course materials focused on accurate pronunciation and enhanced listening skills of both standard French and particularly Quebecan French ‘familiar’ connected speech processes. They consisted of awareness raising activities, the prediction of changes to sentences in citation form once connected speech processes had been applied, dictations, shadowing activities and role plays. Dictation tasks were used as pre- and post-tests to measure participants’ improvements, comprising standard French, familiar Quebec French and combined components. The mean scores increased across the tests for all components, but only results in the standard French and combined components were statistically significant. Again no control group was used in this study.

A potential problem in each of the preceding connected speech processes training studies might be considered to be the participants’ exposure to the post-test words prior to the test. In Brown and Hilferty’s study (1986), the dictation post-test consisted of the words taught during the training program, with only those words scored. As a result, that test measured how well participants had learned the taught words rather than any generalised benefits to listening comprehension, and indeed there was no significant difference between the treatment and control groups in the
listening comprehension test. In the study of Kennedy and Blanchet (2013), the words in the dictation tasks were seen by participants previously when they translated each word into their first languages. No information is given on how closely in time this activity preceded the tests (the pre-test was given during week 3 of the training program, at which time the connected speech processes had been introduced, but not practised). The prior exposure to test content in both studies means that post-test results may not represent real listening gains, and thus raises doubt as to whether the observed improvements would generalise to real speech.

The lack of a control group in the Kennedy and Blanchet (2013) study also means that impact of other factors on the results cannot be totally discounted; such factors would include the progress of time and the greater focus each participant may have placed on their listening skills during the course of the training program.

### 3.6.3 Explicit Training in Phonotactics

As discussed in Chapter 2, Al-jasser (2008) showed, in a study designed specifically to train speech segmentation skills, that participants improved their ability to segment continuous speech after receiving explicit training in English phonotactic constraints and connected speech processes relevant to speech segmentation. The outcomes of this study and the design of the pre- and post-test word-spotting task were discussed in that earlier chapter. Here, the methodology used in the training itself will be examined.

The participants in the training component of the study were 40 native speakers of Qassimi Arabic, majoring in the English language at a university in Saudi Arabia. The total number of participants was divided into an experimental and control group of 20 participants each. Each group was given eight weeks of approximately one and half hours of listening instruction per week. Both groups received listening lessons from
the textbook *Top-Up Listening* 3 (Cooney, Cleary & Holden, 2003), focusing on different features of English connected speech including stress, reduction, contraction, assimilation, and intonation each week. At the end of each week, participants from both groups were told to record 40 minutes of authentic North American speech (from the radio, television or Internet) and transcribe 500 words, highlighting the feature of connected speech that had been the focus of that week’s instruction.

The experimental group also received training in impermissible English phonotactic sequences. The training was conducted in three stages. 1) Awareness Raising: over the first two weeks participants were given background lessons on phonotactics and their usefulness in finding word boundaries. Participants were instructed to memorise a list of twelve illegal syllable-internal consonant clusters. 2) From week three on, in addition to highlighting the feature of connected speech given to both groups, participants in the experimental group were asked to highlight instances of word boundaries consisting of sequences that would be illegal syllable-internally. 3) In-class practice: from week five on, an aural text was played to the experimental group just once and participants were instructed to raise their hands whenever they heard illegal phonotactic sequences. At this point, the tape would be stopped and the participant asked to identify the illegal sequence and where it occurred. This component of the training was designed to increase the automaticity of the identification of sequences.

As discussed in Chapter 2, in contrast to the results achieved after training in connected speech processes material in English (Brown & Hilferty, 1986) and French (Kennedy & Blanchet, 2013), participants who had received only metalinguistic training in connected speech processes in this study did not show any improvement
in the word-spotting post-test. However, the group that had received training in impermissible English phonotactic sequences improved in their detection of target words in the context of those boundary sequences.

Al-jasser did not provide training to raise awareness of impermissible Qassimi Arabic phonotactic sequences that are permissible in English, /bl/, /fr/ and /sw/, nor to provide practice in permitting these to remain unsegmented in word onset instead of treating them as a boundary cue. Perhaps as a result, listeners did not show a difference between the speed or accuracy with which they detected target words in these sequences and in sequences that cue word boundaries in Qassimi Arabic.

In contrast, the word-spotting training developed and evaluated in the present study provided practice in detecting target words in the context of permissible and impermissible English phonotactic sequences and also using word-initial stress cues.

The following section will situate the word-spotting method within current listening teaching approaches.

3.7 Word-Spotting: A Potential New Method

The word-spotting method for training L2 speech segmentation is an activity that clearly sits comfortably within decoding approaches to teaching listening. It is informed by psycholinguistic research into L1 and L2 listening, and the aim of using the method is the development of automaticity in using L2 decoding processes. Like many other activities suggested by writers with a decoding focus, it features controlled exposure to material containing multiple cues to L1 speech details. The task provides practice in the perception of the phonetic details of an utterance to enable recognition of boundary cues, and subsequent speech segmentation and word recognition. As detailed in Chapter 2, phonotactic and prosodic cues appear to be
particularly amenable to acquisition. The focus of the present study is, accordingly, on implicit training using the word-spotting task loaded with prosodic and phonotactic cues to English word boundaries.

The word-spotting method meets specific recommendations of Field (2008b) and Cauldwell (2013), who advocate giving learners the same words in a wide range of contexts and voices. It also meet’s Cauldwell’s recommendation of presenting very short extracts of speech.

Unlike some other approaches to second language listening teaching (see Rost, 2011) it is not interactive or communicative. The task would not be included as an activity in strictly meaning-based approaches. However it does feature authentic materials, conforming to Rost’s idea of ‘genuineness’ in the sense that the nonsense segments could easily form the endings or beginnings of words bordering on the target word. Moreover, there is a continuous transition between the target and nonsense syllable, in contrast to the separation that often occurs in careful, L2-directed speech and listening materials.

The word-spotting method differs from some activities used within the decoding approach as there is no meaning-based instruction beforehand. In each experiment, there was also no explicit instruction beyond an initial information session, which was presented with motivational, rather than instructional goals. In addition the method does not contain a post-listening phase of further activities or repeated practice of difficult material, beyond repeated exposure to the same nonsense-target word string in a later session on another day. However, pre- and post-listening stages could be built around the use of the word-spotting task as a while-listening activity.
3.7.1 Cognitive Validity

To successfully complete the word-spotting task, listeners must segment words from continuous speech contexts. As the processes involved closely resemble those of speech segmentation in real-life listening, the task can be said to have cognitive validity (as per Field, 2011; termed ‘ecological validity’ in McQueen, 1996). Further, as will be detailed in later chapters, performance of the word-spotting task was shown in this study to correspond to performance on other tests investigating segmentation ability.

3.7.2 Benefits of Implicit Training

The mode of training delivery employed in the word-spotting method is similar to that of Cauldwell (2013), in that the task provides implicit training in segmentation skills. Recent second language research has provided evidence that implicit training may lead to more native-like language acquisition than explicit training. In a study on grammatical processing, Morgan–Short, Steinhauer, Sanz and Ullman (2012) found that implicit training in an artificial language evoked L1-like brain activation patterns in high proficiency learners. This effect was not seen in learners trained explicitly. The authors note that this finding provides support for the perspective that language learners fall back on L2 strategies at low proficiency levels, but are capable of achieving native-like grammatical processes and procedural memory with longer exposure and at higher proficiency levels. The results of this study suggest that implicit training may make this possible.

3.7.3 Computer Assisted Training in Decoding Processes

In keeping with the growing trend towards the use of technology to access information and education, a speech segmentation training method might be most beneficial if it could be accessed in a ‘hi-tech’ format. It could be presented online, or
as a CD-ROM, or perhaps as an app for smartphones and tablets, through which participants receive implicit training in a self-study format. This approach corresponds with Vandergrift’s (2004) recommendation that training in speech segmentation be delivered in a language laboratory where listeners can make their own progress towards the “micro-analysis of speech”, and Cauldwell’s (2013) suggestion that training in authentic speech be provided through ‘hi-tech’ methods. Of course, the task could also be one of a battery of small-scale exercises, as suggested by Field (2008b).
Chapter 4

Research Design and Task Validation

4.1 Summary

In this chapter, a set of experiments designed to evaluate the efficacy of the word-spotting task in training L2 speech segmentation is introduced. Stimuli creation for the word-spotting materials is then discussed in depth, followed by a full description of the pilot experiment.

The study as a whole comprised a pilot test and five experiments, one of which (Experiment Two) was discontinued. Experiments One to Four were conducted at a university preparation college, over several weeks each. The earliest experiments were beset by technical difficulties at the college, problems caused by inappropriate post-test materials and increasing attrition as the demands of the preparatory course intensified. The succession of experiments were designed and run as ever more finely tuned attempts to contend with the realities of experimentation outside the laboratory and solve the various problems experienced.

The pilot test was designed to ascertain whether the materials were appropriate for, and the task could be successfully completed by, the target participant group, who were upper intermediate English learners undertaking a university preparation course.
at UWSCollege. Each participant received one session of training in the word-spotting task (50 items).

The pilot-test proved successful and in consequence, Experiment One (Chapter 5) was designed to test the efficacy of the word-spotting task in affecting improvement in a listening comprehension test. The experiment followed a pre-test – treatment – post-test design. Performance in the word-spotting task was compared with that in two other types of listening training: dictation exercises and a connected speech processes training program. Training was held over three eight day sessions of 30 minutes each. While the training task was again performed successfully, the pre-/post-test instrument was found to be too challenging for the participants in this experiment. In addition, technical difficulties at the college meant that the word-spotting task could not be delivered as intended.

In Experiment Two (Chapter 6), a new pre-/post-test was designed and created, but the results of the pre-test showed that this test was again too difficult for the target participant group. Due to the failure of the pre-/post-test to be effective and ongoing technical problems, Experiment Two was abandoned. Note, however, that the training sessions were again well performed.

In Experiments Three and Four (Chapter 7 and 8, respectively), dictation tasks were used as the pre- and post-tests and these were found to be appropriate. In these experiments, training in the word-spotting task was compared with training in connected speech processes, with two groups of participants receiving both types of training over two nine day sessions of 30 minutes each. An additional Response Time (RT) test was also used in Experiment Four in order to ascertain if training in WS led to faster recognition of newly taught words.
The purpose of Experiment Five (Chapter 9) was to conduct the research in a better controlled experimental setting (a laboratory at MARCS Institute) and using a post-test instrument commonly used in psycholinguistic research: an artificial language learning (ALL) task. There were two different participant groups in this study: 1) university students with an L2 background who had lived in an English speaking country for less than two years and 2) university students with a monolingual English background. Participants received three days of 30 minute sessions of either training in the word-spotting task or in connected speech processes, followed by the ALL task on the fourth day.

Across the series of experiments, data was collected on two groups of measures - the performance of participants in pre-tests and subsequent post-tests, and also their performance in the training programs themselves. Performance was measured in the word-spotting task in Experiments Three, Four and Five.

In each experiment, it was hypothesized that if word-spotting training was beneficial, the group receiving that method of training would perform better, or no worse, on the post-test than the groups receiving established methods of training. A significance level of $\leq .05$ was required to accept the hypothesis.

### 4.2 Stimuli

#### 4.2.1 Selection of Training Words

A set of word-spotting stimuli was created that could also be used in future experiments. This stimuli set is different from existing sets of word-spotting stimuli in that it was designed for and recorded using a general Australian English accent and that each stimulus item contains a real word. When the word-spotting task is used in investigative psycholinguistic research, some filler items are included that do not
contain real words, but in its use as a training tool, it is most advantageous if every item contains a real target word and provides exposure to segmentation cues.

4.2.5 Editing

Each recorded stimulus item was extracted using the Praat program (Boersma 2001) and the best token for each speaker of each item was chosen and double-checked by the author's supervisor. Noise surrounding each item's waveform was reduced to zero.

4.2.1.1 Detection of Words without Phonetic Embeddings

It is important that only one real word can be selected from each word-spotting stimulus item. Therefore, a large number of words were required that, when pronounced, do not have any other words embedded in them. Words that can be broken down into more than one possible word could not be included. For example, the spoken word *candidate* contains the embedded words: *can, canned, candid, and, day, date, eight* and *ate*. The presentation of this word to a listener in a word-spotting experiment could result in any of these words being selected.

A computer file version of the Macquarie Dictionary (Butler, 2005) was acquired, which contained Macquarie's own code versions of their phonetic transcriptions. An automated search of this file was conducted using a program written in C# and using Visual Studio 2008 software. The program excluded in its output all words containing a repetition of the phonetic form of another dictionary item, irrespective of spelling; for instance the word *succumb* contains the phonetic embedding *come*. In order to conduct this search, a list of all words of four or fewer letters was automatically created. This list contained 5,426 words. The list was hand-searched

---

8 This program was written by Johnson Chen, programmer at MARCS Institute.
and items to be exempted from the automated search were removed. The full list of excluded words can be found in Appendix A. Through this process, the following categories of words were excluded from the list of potential embeddings:

- very infrequent words in the Australian English context for all but advanced learners (e.g., err /ɜ/). This category includes highly colloquial (gob /gɒb/) and specialised words (e.g., lex /lɛks/, a term in computer science for a program that generates lexical analysers; also Latin for law (Butler, 2005).
- Proper nouns that could not be expected to be spuriously activated in listening, (e.g. Kent /kɛnt/) were also included in this group, most of which were unknown even to the author. – 940 items
- short acronyms (e.g., GST /dʒɪɛsti/). Acronyms that are said as words, for instance COAG /koʊæg/, were also marked for exclusion. – 99 items
- colloquial abbreviations (e.g., comp /kɒmp/, short for competition) – 66 items
- exclamations (e.g. ah /a/) – 40 items
- letters (e.g. D – /di/) – 22 items. P, T, W and Y were not excluded as they form the words ‘pea’, ‘tea’, ‘double, you’ and ‘why’, respectively. U (‘you’) and B (‘be’) were excluded, as they fit into the category of function words, below.
- short function words, including shorter pronouns (e.g. in /ɪn/ and I /aɪ/) and inflectional morphemes (e.g. –er /ə/) – 12 items

Had words such as these remained in the list, there would have been consequences for the automation process. For instance, items such as complex and complexion would have been excluded on the grounds that they contain the colloquial abbreviation comp, which is unlikely to be known by any but the most advanced learners of English. These items and others, such as perplex, would also have been

---

9 The categorization was based on the author’s own usage or understanding of the words’ meanings.
excluded by virtue of the fact that they contain the embedding *lex*, although it is unlikely that any native speaker, let alone a learner of English, would be misled by this item and recognise *lex*.

The assumption that small function words should not invalidate a candidate was made on the basis of Cutler and Norris’ (1988) inclusion of such items, for instance, the word *in* in the stimulus item *mint*, and on the basis of the words’ pervasiveness as embeddings.

3 and 4 phoneme words that clearly contained an embedding of a high frequency word themselves did not need to be marked for exclusion as the automated process would remove these words from the final output, e.g. the infrequent word, *abed* contains *bed*, the acronym, *AJC* contains *sea*, as of course, does the letter *C* itself.

The exclusion from the automation process of the categories of words listed above meant that only the occurrence of legitimate, non-specialized words could invalidate another word in the dictionary. At this point, no value judgment was made on the appropriateness of the word for reasons other than unnecessary invalidation.

**4.2.1.2 Choice of Words for Training Program**

The subsequent automatic search generated 4372 words that did not contain any embedded phonetic words. A decision was then made on which of those 4372 words were suitable for inclusion in the training program and which should be excluded.

It is worth noting here that in the earlier process of exclusion, selected items were made exempt from the list of words that were not permitted to repeat in any other word. In this way, the excluded items were permitted to remain in the files. In contrast, in the following process of exclusion, words were removed from the lists and not permitted to remain in the files. So first, for instance, the word *abb /æb/ was*
excluded from the automation process so that all words containing the sequence /æb/ (such as absent and cab) were kept in the resulting files. At the point described below, however, the word abb was removed from the files as it was not to be included in the final training stimuli.

The primary criterion was that words chosen for inclusion should be ones that upper intermediate level students would be reasonably likely to know. Here the decisions drew on the author’s eight years of experience as an ESL teacher. This was supplemented by the criterion that all words should be listed in materials suitable for advanced second language learners. This was operationalised as listings in a Japanese/English dictionary published for Japanese speakers of English (Shibata, Seiichi & Onishi, 1988), as the dictionary specifies learner levels for each included word.

**General Categories for Exclusion from Training List**

The choice of words to exclude from the training material was informed by both statistical and linguistic considerations. These categories were:

- highly infrequent words (again, for all but exceptionally advanced learners)  
  (e.g., kestrel) – approx. 38%
- names (e.g., Oswald – Lee Harvey Oswald, the man charged with the assassination of JF Kennedy, and Polwarth – a breed of sheep (Butler, 2005), including place names (e.g. Granada) – approx. 27%
- dictionary entries that may have originally contained symbols from an alphabet other than English and appeared with substituted symbols in the file version (e.g., &agrave; bas) – approx. 4%
- all acronyms (e.g. EFL) – approx. 3% of words
- dictionary entries which contained two words (e.g., bishop sleeve) – approx 2%
• highly colloquial words (e.g., boogie) – approx. 1%
• exclamations (e.g. oi) – approx. 1%
• function words (e.g. or), or words that were homophonous with these (e.g., awe) – approx. 1%
• words that would be culturally inappropriate, such as sexual terms and words that could be considered racist or otherwise offensive (e.g. arse) – approx. less than 1%
• letters – 22 items

*Phonologically Based Categories for Exclusion*

Some specifically linguistic considerations were also applied. Words were excluded if they differed only in the manner of articulation of one consonant sound from one part of speech to another. For instance, *elegance* (noun) differs only from *elegant* (adjective) in the consonants /t/ and /s/, which are produced in the same place of articulation (contact of the tongue tip or blade with the alveolar ridge, the bony structure behind the top teeth), but in a different manner (/t/ is produced through the tongue being held against the alveolar ridge, allowing air pressure to build behind the closure, and then promptly released, whereas /s/ is formed through the tongue blade approximating the alveolar ridge creating a narrow channel and resulting in turbulence as the air passes through). Exclusion of these words prevented the misperception of single segments from influencing the error rate.

Words were also excluded when they contained schwa between consonants that could be misperceived as a TR onset forming a word, for example, *referee* and *free*. TR is linguistic shorthand for an obstruent-liquid sequence, as in /tr, fr, fl, bl, kl, and gr/ sequences. Many languages do not allow such sequences and there is evidence for phantom vowels being heard between them (Dupoux, Kakehi, Hirosi, Pallier &
Mehler, 1999). If non-native listeners hypercorrected this error, they could mistake the word *referee* for the word *free* for these reasons. Errors such as these also had the potential of producing an artificially high error rate.

Note that all of the above categories of words were excluded from the training material selection, but were not discarded from the file. They were marked in case of need for future projects. Such a project could, for instance, involve training non-native listeners to perceive reduced vowels, small featural differences or TR sequences accurately for correct word recognition. Further lists of words with particular features were also compiled for potential future projects.

### 4.2.1.3 Assigning Word Frequency

Once the final list of words to be used in the training study had been compiled, the suggested acquisition periods as stated in the Japanese-English dictionary (Shibata et al., 1988) were added. The codes in the dictionary were in fact ‘***’ for those words that are expected to be learned in junior high school, ‘*’ for those learned in senior high school, and ‘†’ for words that are expected to be acquired in the years beyond high school (presumably in university or the workplace), but will be referred to here as 1, 2 and 3, respectively. The words in the dictionary without a rating were the more sophisticated concepts or low frequency words. As mentioned above, words not appearing in the dictionary at all were excluded from the list.

The frequency of each word per million words in spoken English was ascertained from the companion website for a text on word frequency based on the British National Corpus (Leech, Rayson & Wilson, 2001), [http://ucrel.lancs.ac.uk/bncfreq/](http://ucrel.lancs.ac.uk/bncfreq/). The frequency rating given was for the citation form of the word, excluding inflectional variants (e.g., plural, past tense). Where there were words that are homophonous between different parts of speech (for example, ‘back’ can be an adverb or a noun),
the total number of occurrences was recorded, as well as separate frequencies for the phonetic word as it functions in the different word classes.

4.2.1.4 Final list of training words

The resulting list of words to be used in the training study comprised 513 words from the original list of 4372 words. The full list can be found in Appendix B. Note that there are many more words that could be used with native speakers, bilinguals or very advanced learners.

4.2.2 - Formation of Target-Context Strings

4.2.2.1 Nonwords

Nonwords were created through a website program run by the Washington University in St Louis that generates nonwords to the users’ specifications (Balota et al., 2007), http://elexicon.wustl.edu/default.asp. 54 pages of monosyllabic nonwords were generated.

4.2.2.2 Experimental Parameters

The word-spotting materials in this research featured a real word preceding or following a nonsense syllable. Two parameters informed the design of the stimuli sets. As discussed in Chapter 2, two important cues used by native English listeners in segmentation are stress shape and phonotactic sequence (Cutler & Norris 1988; Cutler 2000). As the goal of the research was to train non-native listeners to employ these cues, the main stimulus parameters were thus stress pattern and boundary transition between the real words and nonwords. Hence, nonsense strings were created that would train listeners to make use of the rhythmic patterns typical of English and also to draw attention to phonotactic sequences that suggest word boundaries.
Boundary Transition Conditions

The two boundary transition conditions were Difficult Transitions and Easy Transitions. In the Difficult Transitions set, the nonsense strings were created in such a way that segmentation is difficult due to the syllable contact creating permissible onsets or codas, as in buldagain, oceandelbs, and enscost. That is, ambiguous word boundaries were formed between the real and nonsense words through sequences that could start or end a word. Correct segmentation of these strings requires permissible sequences to be ‘split’. For instance, a legal onset may be split from a following vowel or a legal consonant cluster, such as /nd/, may be split (as when segmenting ocean from the nonsense string oceandelbs). The target words used in this set are vowel initial or start with voiceless stops, /f/, or the sonorants /m, n, l, r, w, y/, thus allowing the creation of cluster onsets and codas. Material in this group was created with reference to the Maximum Onset Principle (Kahn 1976) which states that as many consonants as possible (within the constraints of the sonority hierarchy) will be syllabified into the onset.

The Easy Transitions set features unambiguous word boundaries between the real and nonwords as the sequences are not possible within syllables (e.g., /pt/ or /lr – as in, tap in gerptap and yell in yellrem). Segmentation of these strings should be easier as the boundaries between the nonsense words and real words violate the Sonority Sequencing Principle described in Chapter 2. There were no vowel-initial real words in this set or the creation of illicit vowel-vowel strings, as this may have resulted in the intrusion of the sound [r] appearing between vowels, glottal stops, or other connected speech processes appearing in the recording.
**Stress Condition**

The final factor manipulated in the material was stress placement. In English, stress falls predominantly on the first syllable (Cutler & Carter, 1987). In the *Stress* condition, the real word is monosyllabic or initially stressed and is adjoined to the nonsense context with a reduced vowel (e.g., *bed* in *bedesh* and *thousand* in *nelthousand*). This patterning was chosen not only because of its ubiquity in natural English speech (Gussenhoven & Jacobs, 2005), but because English listeners rely more on contrasting vowel quality across stressed and unstressed syllables than cues from the stressed syllable alone in lexical activation (Cooper et al., 2002). The items in the *Stress* condition also feature the phonotactic parameters discussed above.

**Final Groupings**

The word-spotting stimuli were thus separated into two groups based on the position of the real word in the nonsense string and then each of these two groups was split into three further groups, stress, easy transition and difficult transition. This created six parameter sets. Equal numbers of each difficulty group were put into each set. These were assigned so that the first letter of each was more or less equally spread through the alphabet. The frequency of the target words was also controlled across the sets.

**4.2.2.3 Phonological Considerations**

Many factors needed to be considered in the design of the word-spotting material. For instance, there is a rule in English phonotactics that a lax vowel cannot appear at the end of a word without a following consonant. For this reason, it could be necessary to have tense vowels or diphthongs in preceding nonsense syllables so that the syllable would be phonotactically legal once the real word was segmented from the string. However, Norris, McQueen, Cutler, Butterfield and Kearns (2001) showed
that English listeners’ response times in detecting words bordering lax vowels did not
differ from those bordering tense vowels. In another word-spotting study, Newman
et al. (2011) found that the presence of syllable-final tense or lax vowels in the
nonsense context did not have an effect on segmentation of the string. Therefore, as
well-formedness of consonant-vowel syllable strings doesn’t appear to assist
segmentation in English, it was not seen as needing to be factored into the study.

Other than this, only syllable structures that would occur in real speech were used.
For example, in the Difficult Transition set, a nonsense sequence such as allowbgen
was not used, as successfully segmenting the word allow from the string would result
in a nonsense word with an impermissible consonant cluster /bg/ remaining. A
consonant such as /b/ would also not be found floating unattached to a syllable node
in real speech.

### 4.2.3 Summary on Stimuli Created

A full list of the available training items with their frequency ratings can be found in
Appendix B. A summary of the resulting stimuli, organised by target word position,
follows:

#### 4.2.3.1 Real Word Final

- **Difficult Transition Final (DF):** Correct segmentation of these strings requires
legal syllables to be ‘split’. For instance, a legal onset may be split from a
following vowel, or a legal consonant cluster may be split to form a coda. For
example, ’cost’ in enscost. The real words in this set are vowel initial or start
with voiceless stops, /f/, or the sonorants /m, n, l, r, w, y/, thus allowing the
creation of CC onsets. Material in this group was created with reference to the
Maximum Onset Principle (Kahn 1976).
• **Easy Transition Final (EF):** The context and real word form an illicit boundary; for example, ‘book’ in *fotbook*.

• **Stress Final (SF):** The real word is monosyllabic or has a strong first syllable. The nonsense syllable placed before the real word is a weak syllable, thus forming a WS stress shape; for example, ‘neck’ in *treneck* (‘tre’ contains a schwa). There are no real words with weak first syllables in this set.

### 4.2.3.2 Real Word Initial

• **Difficult transition Initial (DI):** As above; for example, *deep* in ‘deepralts*.

  Vowel final words were bordered by CC onsets to nonsense words. Most of the words in this group have medial CCs.

• **Easy transition Initial (EI):** As above; for example, ‘sock’ in *socknal*.

• **Stress Initial (SI):** As above, but with the weak nonsense syllable being placed after the real word. For example, ‘bed’ in *bedesh* (with ‘desh’ containing a schwa). Again, there are no real words with weak first syllables in this set.

### 4.2.4 Recording

#### 4.2.4.1 Equipment

Prior to recording the stimuli, a program was written in C# and built with Visual Studio 2008 software\(^{10}\) that would transfer the Macquarie code version of the phonetic script into the International Phonetic Alphabet. The Opa 1.0 stimulus presentation software, again written in C# and developed using Visual Studio 2008 software\(^{11}\), was chosen to display the items on the screen during the recording as items can be rejected online and cued to be repeated at the end of the full list.

---

\(^{10}\) The program was written by Johnson Chen, programmer at MARCS Institute.

\(^{11}\) This program was also written by Johnson Chen, programmer at MARCS Institute.
Recordings were made digitally to computer using Cool Edit 2000 software (now Adobe Audition) and an EDIROL UA-25EX USB audio interface. A Rode NT2 condenser microphone was used.

### 4.2.4.2 Speakers

The stimuli were recorded by a number of speakers in order to facilitate generalisation across speakers within the training and beyond. Generalisation across speakers, settings, stimuli contexts and time is a central goal of training programs aiming to affect behavioural or perceptual change, including those found in applied behaviour analysis (Stokes & Baer, 1977), speech pathology (Bernthal & Bankson, 2004) and perceptual training (see Bradlow (2008) for a review).

Due to lower demands on pre-lexical processing, listeners find it easier to identify words when attending to recordings of one speaker than recordings of several. For example, Mullenix, Pisoni and Martin (1989) found that words were more readily recognised in a list spoken by one speaker, than when the same word list was spoken by 15 different speakers. Training the participants using one speaker's voice would make the task easier, but would not assist in generalisation of training effects to real world contexts.

In this research, each word-spotting item was recorded by each speaker. A future training program might feature recordings of the target words said in different nonsense strings by different speakers. This would pose an additional level of difficulty and further facilitate transfer of training outcomes.

As the speakers taking part in the recording would need to read the list of stimuli in phonetic script in order to say them correctly, only speakers who had undergraduate phonetic training were selected. It was also important that the speaker understand
the phonological features being targeted and the speech processes to avoid in the recording. Five speakers were chosen to record the stimuli (three female and two male), all with general Australian English accents.

4.2.4.3 Attention to Phonological Processes

Steps were taken to ensure that the recordings were of a high quality. The speakers repeated each list at least twice. In addition, certain natural English connected speech processes or stress changes were prevented from occurring in the recording. If instances of the following occurred during the recording, the items were rejected and the speaker was asked to repeat those words.

- In English, when the aspirated stops [pʰ] [tʰ] and [kʰ] follow /s/, they become unaspirated, which means that they are produced with less air on their release (see for example, Ladefoged, 2001). This makes these sounds very similar to their ‘voiced’ counterparts, /b/, /d/ and /g/, which are actually voiceless unaspirated stops [p], [t] and [k], respectively. Therefore, placing a nonword ending in /s/ before a real word starting with [pʰ] [tʰ] and [kʰ] is problematic. If recorded as one word, the sequence would cause voiceless aspirated stops to become unaspirated, thus changing the cue. For instance, cost in enscost would sound more like gost. To stop this happening it was ensured that the /s/ did not become ambisyllabic, that is, attach itself to both the nonword it belonged to and the real word following it. Consequently, the stimulus item, enscost, was said as 'ens-kost', with the /s/ clearly belonging to the first syllable, instead of 'enskost' with an ambisyllabic /s/.  

- Medial homorganic consonant clusters, that is, sequences of consonants occurring within a word that are made in the same place of articulation, result in an unreleased first consonant. For instance, in the stimulus item pigkang,
the boundary transition sequence is /gk/. [g] and [k] are both produced by raising the back of the tongue to a position at the rear of the hard palate; the difference between these sounds being voicing. In this circumstance, the [g] would not be released before the [k] is produced. This would create a situation where the real word, pig, was said with an unreleased /g/ ([g’]). This could be a facet to manipulate in later studies, but the decision was made to control for it in this study and ensure that the first consonant in medial homorganic consonant clusters such as these was released.

- A further consideration concerned vowel-initial words in the stress condition in final position, such as lepabsent. There was a risk of liaison occurring in these contexts, with the final consonant of the preceding weak syllable being placed in onset position of the strong syllable (see, for example, Gussenhoven and Jacobs, 2005). In this case, due to the phonological rules of English, the unaspirated [p] of the nonword could become a syllable-initial aspirated [pʰ] preceding the vowel /æ/ of absent. It was necessary to ensure that this did not occur during the recording process.

It must be noted here that such phonetic adjustments in the articulation of the stimuli would result in acoustic-phonetic cues that could signal word boundaries, in addition to the intended stress and phonotactic cues.

The clarity and accuracy of tokens recorded by two of the female speakers was not considered to be of a high enough standard, so the recordings of only three speakers were included in the subsequent studies: JF (female), AW (male), and DL (male).
4.3 Validation

A pilot test was conducted in order to validate the use of the word-spotting task with a learner population (as reported in Cutler & Shanley, 2010)\(^2\). The questions of interest were whether English learners could successfully complete the task and how they responded to the conditions manipulated in the materials.

Ethics approval was granted by Human Ethics at the University of Western Sydney for the pilot test under the title: Micro-Structural Training for Speech Segmentation in a Second Language. The Protocol Number for this project was HREC H6327.

4.3.1 Participants

The pilot test took place at UWSCollege in Westmead, Sydney. UWSCollege is an English language institute affiliated with and wholly owned by the University of Western Sydney. Participants were drawn from the highest level of English language classes at UWS College – English for Academic Purposes, level 3 (EAP III). To gain entry to this course, students must have an International English Language Testing System (better known as IELTS) score of 6 or more, with a minimum score of 5 in all test areas (or equivalent score in another recognized test). An IELTS score of 6 (out of a possible 9) puts students in the category of ‘competent’ English users (see Appendix C).

Participants were 33 upper intermediate students in the EAP III course. The first language of 17 of the participants was Arabic (dialects of the Arabian Peninsula) and the first language of the other 16 participants was Mandarin Chinese. Their average

---

\(^2\) Published under the author’s maiden name.
length of residence in Australia was 20 weeks, with a third of participants residing in Australia less than one month. All participants had normal hearing.

4.3.2 Pilot Test Stimuli

Upon consultation with the staff at UWS College, it was decided to present only the most frequent words in the stimuli sets. 240 audio files were selected for use. These were arranged into four separate lists of 50 word-spotting items, in which the six conditions and two speakers were equally distributed, preceded by ten practice trials. The majority of the words (75%) were monosyllabic (e.g., star, pet), with the remaining words either disyllabic (e.g., perhaps, action), or, in a few cases, trisyllabic (e.g., cigarette, medicine). Only items recorded by JS and AW were used in the pilot study as the recordings of DL were yet to be fully prepared, thus the participants were exposed to both a male and female general Australian English accent. The word-spotting stimuli was programmed for random presentation on DMDX, a Win 32 based display system (Forster & Forster, 2003).

Participants received a questionnaire once they had finished the training (see Appendix D). The questionnaire contained the instruction, “Circle all the words that are true for you. The task was...” Following the instruction there was a list of 15 words or phrases the participants could use to finish the statement. These were: too easy, easy, difficult, too difficult, fun, interesting, boring, helpful, confusing, too long, long, short, too short, tiring, and exciting.

4.3.3 Equipment

The study was conducted using a Lenovo ThinkVantage T61p laptop, running DMDX software, version 3.2.0.1 (Forster & Forster, 2003). A set of Koss headphones was used for audio presentation. Spoken responses were recorded via the microphone input to the laptop.
4.3.4 Procedure

The experimental session consisted of 50 trials preceded by 10 practice trials with feedback. Participants heard one of four lists of word-spotting items which were balanced across conditions (at least seven participants heard each list). Participants were seated at the laptop and wore headphones and the author sat next to each participant. Instructions were given verbally by the author and also appeared on the computer screen. At the beginning of each trial, the words ‘GET READY’ appeared on the laptop screen. The participant then heard a word-spotting stimulus item. The participant responded by saying what they thought was the target word in the nonsense string and the response was recorded by the laptop’s internal microphone. ‘Correct’ and ‘Incorrect’ responses were recorded by the author by either depressing the right or left shift key of the laptop, respectively. The author then pressed the ‘enter’ key to proceed to the next trial. It took participants approximately 5 to 7 minutes to complete the task.

At the end of the experiment, participants were given the questionnaire on their opinions of the task. The author assured participants that their responses were anonymous and that negative opinions were just as highly valued as positive ones as each response could assist in the development of the training program.

Including the time spent giving information to and gaining information from the participants, the whole session took approximately 30 minutes per participant.

4.3.5 Results

It was hypothesised that participants would find the stress final (again, that is, the target word was in final position) and easy transition conditions the easiest to segment and the difficult transition and stress initial conditions the most difficult (EF, EI, SF < DF, DI, SI). It was also hypothesised that the Mandarin and Arabic
participant groups would find the varying phonotactic sequences at the boundary more or less difficult to segment, depending on the phonology of their L1s.

The percentage correct was calculated for each participant over each condition. Error rates per participant showed that all participants scored higher than 50%. The mean percentage correct for the Arabic group was 65%, and for the Mandarin group it was 66%. This difference was not significant.

The effects of the structural parameters manipulated in the stimuli were assessed through a three-way analysis of variance. Word position and condition were the within-subject factors and first language was the between-subjects factor. There was no significant difference between the two language groups on their performance across conditions. There was a significant main effect of word position, with a significantly lower rate of errors when the real word was in final position than in initial position, $F(1,32) = 5.52, p = .03$. There was also a significant main effect of condition, $F(2,64) = 11.24, p < .001$. As expected, words in the easy transition (e.g. socknal, fotbook) condition were significantly easier to detect than those in the difficult transition condition (e.g. deeprals, enscost), $F(1,32) = 14.27, p < .001$, as were words in the stress condition, $p < .001$. Mean scores for all listeners are shown for the six conditions by word position in Figure 4.1.
In addition, there was an effect for the interaction of word position and condition, $F(2,64) = 6.26, p < .004$. An additional within-subjects analysis of variance drew out this effect. The effect of position was significant in the stress condition (trenalck vs. bedesh; $t(32) = 4.66, p < .001$, but not significant in the easy and difficult transition conditions, $t < 1$ in both. Within position, pairwise comparisons showed no significant difference between the conditions with initial target words. However, there was a highly significant difference between the conditions with final target words. Here the stress condition (trenalck) was segmented more easily than both the easy transition (fotbook), $p = .03$, and difficult transition (enscost) conditions, $p < .001$. The easy transition (fotbook) condition was significantly easier to segment than the difficult transition (enscost) condition, $p = .02$. 

Figure 4.1 Percentage of correct word detection, averaged across participants, for the Stress, Easy Transition and Difficult Transition conditions, as a function of target word position in the nonsense item. 

$N=33$
4.3.5.1 Questionnaire

The questionnaire results are shown in Figure 4.2.

![Questionnaire results](image)

**Figure 4.2 Questionnaire results.** The percentage of participants who selected each questionnaire response. N = 33

4.3.6 Discussion

The results of the pilot experiment supported the conclusion that further research using the word-spotting task as a training tool with an L2 population was appropriate.

All the participants answered more than 50% of the items correctly and segmented the material in predictable ways, showing sensitivity to its phonotactic and suprasegmental content. This was not surprising. Recall that the Arabic listeners in Al-jasser’s (2008) study also showed an awareness of the phonological structure of English prior to training.

Participants found it easier to segment the material and detect words in the *Easy Transition* condition, where the phonotactic sequence at the boundary between real and nonwords would be illegal syllable internally (e.g., *fotbook*). They also found
having to split a legitimate syllable internal sequence to place a word boundary
difficult (as in *enscost*). The mean error rate for real words in final position preceded
by a reduced vowel (e.g. *treneck*) was significantly lower than for the other conditions.
Listeners used native English word boundary cues to segment the *Stress Final*
condition, where there was a salient full vowel and accompanying rise in pitch to
compare with a preceding reduced vowel.

However, there were significantly more errors when an initial real word was followed
by an unstressed syllable with a reduced vowel (e.g. *bedesh*). The error rates for the
*Stress* condition with a final target word were very similar to those for the *Difficult*
*Transition* (*deepralts* and *enscost*) sets. It is highly probable that this is because the
initial real words followed by unstressed syllables have similarities to the *Difficult*
*Transition* condition. It is likely that participants had trouble knowing where the
word ends. Given the ubiquity of function words in streams of continuous English
speech, this is clearly a common problem in authentic listening and an area which
may benefit from repeated practice in segmenting such material.

The results of the questionnaire were very positive, with most participants rating the
test as interesting, helpful and exciting. None of the participants rated the test as
being too easy, too difficult, too long or boring. Responses showed that they found
the task challenging, but not too demanding to allow successful performance.

Due to the success of the validation of the task, the use of the word spotting paradigm
to train participants to use native English speech segmentation strategies was deemed
feasible.

4.3.6.1 Impacts on Research Methodology

It was initially hoped that an analysis of participant responses would result in a list of
phonotactic sequences to target in the subsequent training, tailored to Arabic and
Mandarin learner errors respectively. However, a survey of the error rates versus correct responses for different phonotactic sequences showed no consistent pattern, showing that the phonotactic contents of the sequences were not a significant factor in performance. Indeed, there was no significant difference between the two language groups on any result. It was clearly the structure of the stimuli rather than the listeners’ first languages that caused significant effects in this pilot study.

Due to the lack of difference in performance of the two (structurally very different) language groups, the initial plan to modify the materials to target a specific L1 was abandoned for the decision to train a variety of L1s with the full training sets of the first two frequency levels. This was considered a more practical decision given the fluctuation in student enrolments from different countries across teaching terms at UWSCollege, and consequently, the strong likelihood of low participant numbers. In addition, as the material design was based on the structure of English, and not on any particular target L1, the training was deemed to be suitable for all L1 groups.

UWSCollege has a diverse student population. The participants in the following experiments were from a range of L1 backgrounds: Arabic, Bengali, Cantonese, Cebuano, Dutch, Farsi, German, Hindi, Hokkien, Japanese, Korean, Mandarin (and the Southwestern Mandarin dialect), Pashto, Punjabi, Russian, Spanish (Mexican and Peruvian dialects), Tagalog, Telugu, Turkish, Urdu, Vietnamese, and Wu (in addition to the Shanghainese and Hangzhounese dialects of Wu).

As a consequence of the disparity in the way rhythmic and phonotactic cues signal word boundaries across these languages, analysing the word-spotting boundary conditions separately may have led to L1 effects. The impact of the L1 on the performance of the word-spotting task has already received much attention (as
discussed at length in Chapter 2) and is not the focus of the present study; the goal of which is to examine the efficacy of the word-spotting task as a method of training English word boundary cues to people from a range of language backgrounds. Performance at the word-spotting task was not examined in Experiment One or Two. In Experiments Three, Four and Five, the word-spotting data was collapsed across conditions to provide an overview of cross-linguistic training effects.
Chapter 5

Experiment One

5.1 Aim and Design

The aim of Experiment One was to investigate whether training in the word-spotting task would improve speech segmentation. To this end, Experiment One sought to establish if the word-spotting method is better than, or on par with, variations on existing listening training programs.

As described in Chapter 3, while not directly aimed at teaching speech segmentation skills, some studies have investigated the teaching of listening skills through the training of decoding processes. These studies employed dictation (Kiany & Shiramiry, 2002; Matthews & O’Toole, 2013) and training in connected speech processes (Brown & Hilferty, 1986; Kennedy & Blanchet, 2013). In incorporating elements of these training methods to train the control groups, not only could the efficacy of the word-spotting method be evaluated, but a further assessment of the effect of dictation and training in connected speech processes on listening comprehension could be provided.

The experiment followed a pre-test – treatment – post-test sequence and was based on a between-subjects factorial design with the dependent variable being scores on a
listening comprehension test. The total number of participants was divided into three groups and received each of three different listening training programs in turn: 1) the word-spotting method, 2) dictation exercises and 3) a connected speech training program. It was thought that withholding potentially beneficial training from any one group would be unethical. This format followed a Latin square design, which is often used in medical research to compare the effects of different drugs or treatments (Preece, 1991). Training was held over three eight day sessions of 30 minutes each, with a post-test held at the completion of each session.

5.2 Hypotheses

1. Participants who received training in dictation and connected speech processes will show improvement in listening comprehension scores, in keeping with empirical evidence of their effectiveness.

2. Repeated exposure to word-spotting training will result in improved scores on a listening comprehension test, on par with or higher than scores received by participants in the control training groups.

5.3 Method

5.3.1 Participants

Participants were 21 self-selected students in the highest level of the English for Academic Purposes course (at that time, EAP III) at UWSCollege. They received payment for their involvement in the course at a rate of $25 at the completion of each of the three training sessions or a total of $80 if each day of each session was completed. If participants missed a training day in any session, they would receive $5 for each hour completed in that session.
The participants’ language backgrounds were grouped as follows for the purpose of analysis: 6 Chinese, 6 Indo-Iranian, 5 Japanese, 3 Vietnamese and 1 Tagalog. The Chinese language group can be further classified into dialect groups: 3 Mandarin, 2 Cantonese and 1 Wu speaker. The Indo-Iranian language group comprised 3 Farsi, 2 Urdu and 1 Bengali speaker. 16 of the participants were female and 5 were male. The participants had an average length of residence in Australia of 11.5 weeks, with two-thirds of them having lived in Australia less than eight weeks.

32 students completed the pre-test and the total number was divided and randomly assigned into three groups. Six students did not return after the pre-test and another five withdrew over the first days of training. The numbers, language background and sex of the participants who completed the training are reported in Table 5.1 for each group. The acronyms in the table refer to the order of training treatments; for instance, Group WCD received word-spotting training in session 1, training in connected speech processes in session 2, and dictation training in Session 3.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Language</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCD</td>
<td>8</td>
<td>Chinese – 3</td>
<td>Female – 8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vietnamese – 2</td>
<td>Male - 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indo-Iranian – 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Japanese – 1</td>
<td></td>
</tr>
<tr>
<td>CDW</td>
<td>6</td>
<td>Indo-Iranian – 3</td>
<td>Female – 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Japanese – 1</td>
<td>Male - 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tagalog – 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vietnamese – 1</td>
<td></td>
</tr>
<tr>
<td>DWC</td>
<td>7</td>
<td>Chinese – 3</td>
<td>Female – 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Japanese – 3</td>
<td>Male - 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indo-Iranian – 1</td>
<td></td>
</tr>
</tbody>
</table>

*Table 5.1 The composition of the three training groups.*
5.3.2 Stimuli

5.3.2.1 Training Programs

Although the content of the training differed, particular aspects were held constant between the programs. First, each training program was delivered by the author in the same classroom using the same equipment. Second, the materials were all recorded using the author’s voice, alongside the voices of other speakers. Finally, feedback was featured and played an important role in each training program.

Word-Spotting

The design and creation of the word-spotting materials were discussed in Chapter 4. An extra feedback screen was written into the DMDX program for this experiment.

As 47% of the participants in the pilot experiment had rated that training format as ‘short’, and 9% as ‘too short’, while only 6% had rated it ‘long’ and no participants had rated it as ‘too long’ (see Chapter 4), the number of training items presented was doubled in Experiment One. In addition, because each training session in Experiment One was to run for eight 100 item sessions, as opposed to the one 50 item session in the pilot experiment, all stimuli available in the highest and second highest frequency groups were used. A full list of the training items used can be found in Appendix B.

The material was divided into three sets of three as follows: A01, A02, A03, B01, B02, B03, C01, C02, C03. The set groups A, B and C each contained one instance of a stimulus item, but in each set the item was recorded by a different speaker. For instance, the item kebig (from the stress final condition) appeared in each set group, but in set group A it was recorded by AW, in B it was recorded by DL and in C by JF. All of the available items appeared in the three files in set A and were repeated in the
three files of set B and again in set C. The repetition of items would reinforce what different words and, most importantly, boundary sequences sound like when said by different speakers. Seeing the full ‘context-target’ string written on the screen after each attempt might have a similar effect in improving L2 listening as the use of L2 subtitles in movies (see Mitterer & McQueen, 2009). A helpful sound-form association of the boundary would be created through seeing the written feedback and further reinforced with each reoccurrence of the item paired with its feedback.

As usual, the task given to the participant upon hearing each item was to identify the target word in the string. However, unlike the pilot experiment and versions of this task used in psycholinguistic research (McQueen, 1996) where responses are spoken, in this training version the participants were asked to type their responses. This change was made due to the fact that many participants would be receiving the training at the same time in a classroom setting and the resulting recording quality may be poor. As in the pilot experiment, there was no measure of reaction time.

**Dictation**

The dictation exercises were chosen from two sources: *Adventures of an Australian English Teacher: English Listening Practice* (Salter, 2003) and *Listening to Australia: Intermediate* (Brawn, 2002). Both are ESL textbooks that focus on Australian English and feature listening activities recorded in the Australian English accent. Australian English was used so that accent would not constitute a confounding variable between the training programs and to provide maximum benefit to the students participating in the study.

---

13 Some of DL’s recordings were deemed not to be of high enough quality to include in the training and in those cases only two examples of the item appeared.
As each word-spotting training set contained 100 words, it was decided that each
dictation should also contain approximately 100 words. Aural texts were chosen from
each book on the basis of the frequency of the vocabulary contained therein and the
ability to find an appropriate 100 word portion. Nine texts were chosen in total: 5
from Brawn (2002) and 4 from Salter (2003). These can be found in Appendix E.

Both textbooks included recordings on CD. The recordings by Salter (2003) were
found to contain instances of unnatural intonation, speech rate and pauses within
clauses, so the author recorded herself reading two of the texts and replaced the
originals with these. This was advantageous as the other training programs also
featured the author’s voice alongside those of others. These recordings were made
digitally to computer using Cool Edit 2000 software and an EDIROL UA-25EX USB
audio interface, using a Rode NT2 condenser microphone.

Participants were provided with double-sided, lined, A4 sheets of paper on which to
write their dictations and transcripts of each of the aural texts upon the completion of
each dictation exercise.

**Connected Speech Processes**

The textbook used to conduct the training in connected speech processes was *Sound
Advice: A basis for listening*, (Hagen, 2000). This text was chosen from Brown’s (2006)
review of nine ESL textbooks that cover instruction in connected speech processes.
Of these nine, Brown most highly recommends the Hagen text, though
acknowledging that this may be due to his own “residual, and perhaps old-fashioned,
need to teach things explicitly, rather than to rely on student discovery learning” (p.
33). As one aim of the study was to compare the efficacy of explicit and implicit
training, the teacher-centred focus of the textbook was considered particularly
appropriate.
The first three chapters of Hagen (2000) were chosen for use in the experiment.

*Chapter One - Reductions covered:*

- word stress
- reduced vowels
- reductions in function words
- reductions in form across words (e.g., ‘going to’ to ‘gonna’).

*Chapter Two - Introduction to Linking contained lessons on:*

- linking across vowel-initial words in connected speech (e.g., ‘How much is it?’ is said as ‘how muh chih zit?’)
- linking of identical consonants (e.g., ‘been no’ is said as ‘beeno’ rather than with two distinct /n/ sounds).

*Chapter Three - Special Sound Changes covered:*

- flaps (e.g., ‘butter’ said with a flapped /t/ ([ɾ]) sounds to the unfamiliar ear like ‘buda’)
- glottal stops (a sound that appears often in English connected speech, but is most commonly associated with the Cockney English accent, as in *bu’a* [bʌʔə] for ‘butter’)
- -nt reduction (the omission of the /t/ and nasalised flapping of /n/ as may be found in *twenny* for ‘twenty’)
- the reduction of ‘of’ to [ə] (as in *I want some-a that*)

---

14 It also contained lessons on ellipsis, but the author has found that students at the upper intermediate level generally have a good knowledge of these forms.
The textbook was only available with CDs recorded in the standard American English accent. In order to keep accent consistent across the training groups, the exercises were re-recorded in the general Australian English accent by the author and a male colleague. Strict attention was paid to the quality of the recordings in order to ensure that the connected speech processes being taught in each exercise were present in the speech. Again, the recordings were made digitally on computer through Cool Edit 2000 software using a Rode NT2 condenser microphone and an EDIROL UA-25EX USB audio interface.

Participants were given photocopies of the relevant pages from Hagen (2000) at the start of each topic.

**Equivalency of Training Programs**

Equivalency of training input and quality was an important concern. Two particular concerns were that, in comparison to the word-spotting and dictation training programs, the teacher-delivered training in connected speech processes would: 1) create a perception of greater teacher support and interest in their learning, and 2) provide more speech input.

To address this issue, the LENA System was trialled in the first few days of Experiment One. The LENA System is a device that can be worn around the neck. It is sensitive to speech input and generates an estimate of the number of words spoken. Unfortunately, the LENA System was found to be impractical in this context. The data needed to be uploaded after each training class, which led to delays in commencement of subsequent classes. Such delays were considered undesirable as they increased the amount of time participants needed to wait at the college and therefore also increased the risk of attrition.
In order to compensate for any additional exposure to ‘teacher-talk’ in the connected speech processes training program, the author ensured she spent time speaking to participants in the dictation and word-spotting programs before and after the training sessions and took an interest in them as individuals.

5.3.2.3 Pre- and Post-tests
As discussed earlier, the most effective instrument with which to study speech segmentation is the word-spotting task. Because the word-spotting task was the experimental treatment here, clearly a different instrument was required to measure improvement in segmentation skills. Improvements in speech segmentation could logically be expected to have flow on effects for listening comprehension, so a listening comprehension test was chosen as the post-test.

The listening component of the Cambridge Certificate of Advanced English (CAE) was chosen as the pre- and post-tests based on recommendations in Assessing Listening (Buck, 2001), which provides an overview of the research into listening comprehension assessment and an evaluation of professionally made listening examinations with high construct validity (that is, the tests have been shown to measure what they purport to measure).

Buck (2001) discussed the Cambridge First Certificate in English (FCE). This test is at the intermediate level and may have been too easy for our upper-intermediate participants. The higher level CAE examination was thought to be appropriate for the participant group as students are advised to have an IELTS 6.5 or higher to sit for it and our participants were required to have an IELTS of 6 or more to enter the EAP III course at UWSCollege. In addition, the certificate is said to offer “a high-level qualification for people who want to use English for business and study purposes”
This was deemed appropriate for this study’s participants, who were enrolled in university preparatory courses.

The particular text used was *CAE Practice Tests Plus* (Kenny & Newbrook, 2008). It contains six past examinations of the Cambridge Certificate in Advanced English, four of which were selected for use in the experiment. In this experiment (and all following) different versions of testing instruments were used in pre- and post-tests. The rationale was that pre-university admission students would remember some contents from test to test, and might also confer with others after the tests to gauge their performance.

The materials in the listening component were recorded in a variety of British English accents. The professional ESL examination most often used in Australia and recorded in the Australian English accent is IELTS. However, many of the participants would have already encountered the practice materials in their own preparation for the IELTS test. As each training group would be exposed to the British English accents in the tests, none would be at a comparative disadvantage.

Each CAE practice test in Kenny and Newbrook (2008) follows a set structure. Each is 40 minutes long, has four parts and 30 questions in total. Listeners are guided through the test by verbal instructions at the start of each part. The initial instructions provide context for the listening activity, for instance, “You hear two guests on a program discussing travel and holidays.” Part 1 consists of three short verbal extracts with two multiple choice questions on each. Part 2 features a long aural text accompanied by a sentence completion task. The sentences paraphrase rather than echo the text. Part 3 consists of a long aural text with multiple choice questions. Part 4 comprises five monologues on a common theme and a task in which multiple items are chosen from a list. Some of the earlier practice tests in
Kenny and Newbrook (2008) include study tips, but these were not included when the tests were copied for the participants’ use.

5.3.2.4 Questionnaire

At the end of the experiment, participants were given a questionnaire. On the first page of the questionnaire, three questions were asked: We compared three types of training that were intended to help you listen to speech. 1) Which training type was most helpful?; 2a) Which training type was the least helpful?; 2b) How do you think this training could be improved?

The three subsequent pages were the same as the questionnaire given to participants in the pilot experiment, with the inclusion of pages referring to dictation and connected speech training. Please see Appendix F for the full text of this questionnaire.

5.3.3 Procedure

The total number of participants was initially randomly divided into three groups of 10 participants each. All three of the groups received all three of the training treatments, but in different orders. The groups will be referred to here as WCD, DWC and CDW, according to the order in which they received the three training methods. For instance, WCD received word-spotting training in the first two-week session, training in connected speech processes in the second and dictation in the third. The training schedule is shown in Table 5.2.

Training took place at the end of the day’s courses at UWSCollege, from 2:10 to 3:40 in the afternoon. The training programs were run by the author, consecutively. Attendance times for each group were rotated every day in order to avoid the confounding variable (for example, potential for higher drop-out rates) of participants
in one training program always having to wait for the last session in the afternoon. Participants received a complete timetable for their group at the start of the experiment.

<table>
<thead>
<tr>
<th></th>
<th>Group WCD</th>
<th>Group DWC</th>
<th>Group CDW</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Listening Comprehension Pre-Test</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Weeks 1 &amp; 2</strong></td>
<td>Word-Spotting</td>
<td>Dictation</td>
<td>CSP</td>
</tr>
<tr>
<td><strong>Listening Comprehension Post-Test 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Weeks 3 &amp; 4</strong></td>
<td>CSP</td>
<td>Word-Spotting</td>
<td>Dictation</td>
</tr>
<tr>
<td><strong>Listening Comprehension Post-Test 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Weeks 5 &amp; 6</strong></td>
<td>Dictation</td>
<td>CSP</td>
<td>Word-Spotting</td>
</tr>
<tr>
<td><strong>Listening Comprehension Post-Test 3</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2 Training schedule for Experiment One. Training in connected Speech processes is referred to here as CSP.

5.3.3.1 Pre-test

Participants were given the pre-test preceding week 1 while seated together in the same classroom. They received photocopies of CAE listening practice test 2 (Kenny & Newbrook, 2008). The CD was played on a Sanyo MCD-Zx530F Portable CD Player.

At the start of each part of the test, participants were given information on what they would hear and complete instructions on how they were to perform the tasks. Their understanding was confirmed by the author before proceeding. Participants were given time to look at the questions before each part started. They heard a tone at the start of each piece and heard each twice.

5.3.3.2 Word-spotting

On the first day of the word-spotting training, participants were given a rationale for the training and a chance to practice the task. Due to technical difficulties at
UWSCollege, the word-spotting training was not conducted as individual, computer-based training. Instead, the training was delivered to the classroom as a whole using DMDX software on a Lenovo ThinkVantage T61p laptop and through Panasonic SB-PM21 speakers. The laptop was placed on the teacher’s desk, facing toward the participants. The font of the DMDX program was increased to 26, so it could easily be seen by all participants, who sat close to the laptop. The author sat beside the laptop and controlled progress through the program, moving to the feedback screen when all participants had finished writing their answers and responding to any requests to slow down or increase the speed of movement between trials. The participants were given sheets of paper with numbered lines on which to record their responses. When each feedback screen was displayed, the participants marked their answers as correct or incorrect. At the end of each day’s training session, the participants tallied their scores for their own interest and motivation. The answer sheets were collected by the author at the end of each training class and re-tallied by the author during the final analysis.

5.3.3.3 Dictation

The training program was run by the author. Participants were first given double-sided, lined, A4 sheets of paper on which to write their dictations. The procedure used followed Kiany and Shiramiry’s (2002) study which provided empirical evidence for the use of dictation activities in improving listening comprehension. Participants were first told the topic and title of the text so that they could predict what they would hear in it and activate a conceptual framework in advance. If the aural text contained a dialogue, the names of the speakers were written on the board, so that participants could better organise their scripts.
The participants then heard the passage in its entirety. Audio files were played using iTunes software on a Lenovo ThinkVantage T61p laptop, through Panasonic SB-PM21 speakers. The aural text was first played in its entirety and then played again, but paused after each “meaningful chunk”, with the participants being given the chance to write down what they had heard. The whole passage was then played again so that the participants could check what they had written and make any necessary changes. The participants were subsequently given the typed script and the aural text was again played in its entirety, in order to provide the participants feedback on their performance. A discussion was held afterwards so that participants could ask questions about words they were unfamiliar with. Participants also had the opportunity to comment on difficult aspects of the passage and the instructor gave further examples of such structures when possible. The dictations were not assessed and were therefore not collected from participants.

5.3.3.4 Connected Speech Processes

Participants were first given a photocopy of the chapter to be taught from Hagen (2000). The lessons in each chapter were taught explicitly by the author and participants were welcomed to ask questions. The recorded exercises were played using Windows Media Player software on a Lenovo ThinkVantage T61p laptop, through Panasonic SB-PM21 speakers.

Once the participants had completed the exercise they were asked to compare their answers with the person sitting next to them. The author asked the pairs if their answers were the same and encouraged discussion on what the correct answer may be and why. This technique stimulates interest in material and encourages greater attention to listening when recordings are played subsequently. The recording was
then played a second time so that participants could check their answers. The correct answers were then given by the author.

5.3.3.5 Post-tests and Questionnaire

The post-tests followed the same procedure as the pre-test. Post-test 1 was Practice Test 3, post-test 2 was Practice Test 4, and post-test 3 was Practice Test 5 from Kenny and Newbrook (2008). After post-test 3 - the final post-test - the participants were given the questionnaire and told their responses would remain anonymous.

5.3.3.6 Number of Participants Completing Each Session.

Table 5.3 shows the percentage of training in each session completed by the number of participants in each group.

<table>
<thead>
<tr>
<th>Word-Spotting</th>
<th>CSP</th>
<th>Dictation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCD</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>75 - 99%</td>
<td>100%</td>
<td>75 - 99%</td>
</tr>
<tr>
<td>50 - 74%</td>
<td>50 - 74%</td>
<td>50 - 74%</td>
</tr>
<tr>
<td>0 - 49%</td>
<td>0 - 49%</td>
<td>0 - 49%</td>
</tr>
<tr>
<td>DWC</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>75 - 99%</td>
<td>75 - 99%</td>
<td>75 - 99%</td>
</tr>
<tr>
<td>50 - 74%</td>
<td>50 - 74%</td>
<td>50 - 74%</td>
</tr>
<tr>
<td>0 - 49%</td>
<td>0 - 49%</td>
<td>0 - 49%</td>
</tr>
<tr>
<td>CDW</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>75 - 99%</td>
<td>75 - 99%</td>
<td>75 - 99%</td>
</tr>
<tr>
<td>50 - 74%</td>
<td>50 - 74%</td>
<td>50 - 74%</td>
</tr>
<tr>
<td>0 - 49%</td>
<td>0 - 49%</td>
<td>0 - 49%</td>
</tr>
</tbody>
</table>

*Table 5.3 The percentage of each training session completed by the number of participants in each training group*
5.4 Results

21 participants commenced the training, but although training attendance rates were high (20 participants completed all three programs) only 14 completed all three post-tests.

Unless otherwise specified, all analyses were conducted using a repeated measures analysis of variance (ANOVA), with training group as the between-groups factor and an alpha level of .05. Assumptions of normality and homogeneity were satisfactory unless otherwise discussed.

Only significant main effects and interactions are discussed within this chapter. All other results from Experiment One can be found in Appendix G.

5.4.1 Pre-Test

Participants performed poorly on the pre-test, with only three participants gaining scores higher than 50%, and all three of those participants scoring just 53%.

Individual pre-test scores are shown in Figure 5.1.

![Figure 5.1 Pre-test scores. N=21](image-url)
Three bands of scores can be seen in the data. The three participants who narrowly passed the test received scores substantially higher than the rest of the group (1.87 above the standard deviation of 3.19). The scores for these participants can be seen in Figure 5.1 under the IDs 7, 9 and 12. There is also a set of four participants who received considerably lower scores, more than 1.27 below the standard deviation (participants 3, 5, 10 and 21). Two-thirds of the participants are normally distributed in the middle. These scores were grouped as High, Mid and Low scoring participants for future analysis under the variable ‘Results Groups’.

A univariate analysis of variance revealed no significant difference in pre-test scores between training groups, first languages, length of residence, or sex. The mean score across the three groups was 33.48%. (The mean score for WCD was 30.88%, for DWC it was 38% and for CDW it was 31.67%.

5.4.2 Pre-test - Post-Test 1

All 21 participants attended post-test 1. The mean scores of all participants combined improved from the pre-test to post-test 1, \( F(1, 20) = 5.65, p = .03 \). The mean score across all training groups was 37% (up from 33.48% in the pre-test). The mean scores for each training group from the pre-test to post-test 1 are shown in Figure 5.2. These show that the groups that received training in dictation and connected speech processes improved over the two time conditions, while the mean score of the group receiving word-spotting training did not change. However, the main effect of ‘training group’ and the interaction between ‘training group’ and ‘test number’ was not significant.
5.4.3 Pre-Test - Post-Test 2

A total of 18 participants completed post-test 2: 5 in WCD, 7 in DWC and 6 in CDW. The participants found post-test 2 difficult. The mean score across all groups was just 29%. Each training group attained a lower mean score on the second post-test than in the pre-test or post-test 1. There was a significant effect of ‘test number’, $F(1, 17) = 6.62, p = .02$.

The scores of the group that received word-spotting in session 2 decreased less than those for the other two training groups, but there was no ‘test number’ by ‘training group’ interaction. Results for each group across the pre-test and post-test 2 are shown in Figure 5.4.
Scores from the three different results groups (High, Mid and Low) were analysed according to their pre-test scores across the pretest and post-test 2. There were no significant main effects or interactions.

5.4.4 Pre-Test - Post-Test 3

A total of 16 participants completed post-test 3; 5 in WCD, 7 in DWC and 4 in CDW. The mean score across all groups was 30%. The group that received word-spotting training in session 3 was the only group that displayed an increase in mean score between the pre-test and post-test 3. In this analysis, neither the main effects of ‘test number’ or ‘training group’, or the ‘test number’ by ‘training group’ interaction were significant. The results are shown in Figure 5.5.
5.4.5 Pre-Test and All Post-Tests

A total of 14 participants completed the pre-test and all three post-tests (3 in WCD, 7 in DWC and 4 in CDW). Across the four tests there was a significant effect of test number, $F(3, 13) = 5.47, p = .004$. Scores in post-test 1 were 7.78% higher than in post-test 2 ($p = .04$), and 8.94% higher than in post-test 3 ($p = .02$). However, there was no significant main effect of ‘training group’ or interaction of ‘test number’ by ‘training group’. Figure 5.6 shows the results for those participants who completed each post-test across each training group.
5.4.6 Questionnaire

The 14 participants who sat the final post-test were given the questionnaire on its completion. As can be seen in Figure 5.7, the participants rated the training in connected speech processes most favourably on the question of how helpful they viewed each training program to be.

![Figure 5.5 The mean percentage of correct scores for each training group across the pre-test and all post-tests. N=14](image)

![Figure 5.6 Participant responses to the questions, “Which (training) type was the MOST helpful?” and “Which type was LEAST helpful?” N=14](image)
The other responses can be divided into positive, neutral, negative and very negative evaluations. The number of participants which selected each response for each training program is displayed in Table 5.4.

<table>
<thead>
<tr>
<th></th>
<th>Word-Spotting</th>
<th>Dictation</th>
<th>CSP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Positive</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fun</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>interesting</td>
<td>13</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>helpful</td>
<td>12</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>exciting</td>
<td>10</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>39</td>
<td>44</td>
<td>43</td>
</tr>
<tr>
<td><strong>Neutral</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>easy</td>
<td>4</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>short</td>
<td>2</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>6</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>difficult</td>
<td>4</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>confusing</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>long</td>
<td>2</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>tiring</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>11</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td><strong>Very negative</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>too easy</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>too difficult</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>boring</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>too long</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>too short</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>8</td>
<td>8</td>
<td>2</td>
</tr>
</tbody>
</table>

*Table 5.4 Questionnaire results. The number of participants (total = 14) who selected each response for each training program*

The questionnaire results show that the connected speech processes training program was considered the most helpful in response to the direct question, “Which type was the MOST helpful?” Interestingly, the dictation training was the only training type that 100% of participants nominated as ‘helpful’ in the second part of the questionnaire. The word-spotting and dictation training were considered less helpful by a greater number of students.
In the second section of the questionnaire, the word-spotting training received the least favourable ratings and the most unfavourable ratings in the questionnaire, while the training in connected speech processes received the most favourable ratings and the least unfavorable ratings.

The preference for the connected speech processes training over the other two training programs may in some respects reflect a cultural bias towards explicit, teacher-led instruction over implicit, student-discovery programs.

5.5 Discussion

Despite having achieved IELTS scores that were the equivalent of the standard necessary to sit the Certificate of Advanced English examination, the listening component of the CAE was clearly too difficult for the participants in this experiment. Very few participants achieved scores higher than 50% in each post-test. The tests were also clearly of different difficulty levels, with most participants performing far better in the pre-test and post-test 1 than in post-tests 2 and 3.

A potential factor in the difficulty of the tests for the participants in this study is that each section of the tests required listeners to make use of contextual knowledge to answer questions correctly. The participants were not trained in exploiting context in this experiment. Accordingly, in retrospect, the test was not an appropriate measure of the skills that were trained.

The large variety of British accents of the speakers may have also posed a challenge for the participants. Another factor worth considering is that the room in which the tests were taken had a very high, cathedral ceiling, which caused perceptible sound reflection. This may have made the test recordings more difficult to understand and have impacted on results.
In post-test 1, the group that received word-spotting training in the first session performed poorly compared to the groups that received training in dictation and connected speech processes. The scores for those groups increased from pre-test to post-test 1, while the scores for the word-spotting group remained the same across tests. However, the post-test 2 scores for the group that received word-spotting training in session 2 decreased less than those for the other two training groups. Furthermore, in post-test 3, the group that had just received word-spotting training was the only group that displayed an improvement in performance.

Due to the complexity of the post-tests and resulting poor participant performance, it is impossible to make categorical statements about the outcomes of this experiment, but there are indications are that the word-spotting training is at least as effective as established L2 listening training methods.
Chapter 6

Experiment Two: Pilot Test

6.1 Aim

Due to the inappropriateness of the post-test materials used in Experiment One, a search commenced for new styles of pre- and post-test. A pilot test was run to assess the suitability of the new tests.

A further aim of the pilot test was to ensure the technical issues experienced with the UWSCollege computers in Experiment One had been resolved.

6.2 Background and Rationale for the New Tests

Construct validity is an important concept in language testing (Buck, 2001; Hughes, 2003). Features of language ability can be said to be explained by sets of theoretical constructs. A test should correctly and directly test the theoretical constructs of interest, without testing extraneous abilities – to do so would add construct-irrelevant variance to the testing process and outcomes (Buck 2001).

In this case, the theoretical construct of interest is speech segmentation. At this point, new tests were needed which would directly test speech segmentation ability, without also testing higher level processes, such as inference. In reality, it is hard to separate one language ability out completely from others. However, there are some
tests that allow for more direct assessment of segmentation ability than others and in which construct-irrelevant variance can be minimised.

6.2.1 Listening Cloze Test

A listening cloze test was designed and created. The cloze test was first described (and the name coined from closure in Gestalt Theory), by Taylor (1953). Taylor examined its use as a test of text readability and published his findings in a journalism journal. He also noted that the procedure could be used to compare the reading abilities of individuals.

In a traditional cloze test, sentences or passages are presented in written form, with some words omitted (either every nth word or randomly assigned) and replaced with blank spaces where the reader may insert their answers. The traditional cloze procedure tests reading ability, vocabulary knowledge and general language proficiency. In listening cloze tests, aural texts are played to the listeners, who complete the gaps in the written text with the words heard.

A number of studies (Gregory-Panopoulos, 1966; Kaplan & Jones, 1970; Templeton, 1977; Williams, 1962) have found that listening cloze tests in which every nth word is deleted measure listening proficiency with more reliability than traditional tests of listening comprehension such as multiple choice tests.

When used as a listening test, often the only difference from a traditional cloze test is that a recording of the passage is played. The whole passage or sentence is presented for the listener in writing on the answer sheet, with gaps for the intended words.

This was unsatisfactory for the purposes of this study as the segmentation would have already been done for the listener and the listening activity reduced to an exercise in single word recognition. In addition, Buck (1988) states that some listeners treat
these types of tests as traditional cloze tests and insert a word they have guessed from whatever context has been provided without attending to the aural text.

Instead, it was decided to base the listening cloze test on a technique developed by Templeton (1977). In Templeton’s listening cloze test, words are replaced with what he referred to as ‘electronic bleeps’. Templeton used two, roughly three and a half-minute long university-lecture style aural texts, in which bleeps where inserted over every 15th word. The test results were compared with teacher and peer ratings of each listener’s listening proficiency and reliability was found to be high.

Given the aims of the study, it was necessary to try to separate general listening or language proficiency from speech segmentation ability. In this study, unrelated sentences were used rather than passages, to avoid words being predictable due to context.

Each sentence was played once in its entirety and then a second time, in which one word in each sentence was replaced with a bleep. Again, to avoid predictability and also to ensure each word of the sentence was attended to, the procedure of choosing every n-th word was not followed. The choice of word to be replaced was made through rotation through parts of speech. This was done in order to prevent potential researcher bias in choosing words that were thought to be more difficult (or easier) to segment.

The answer sheet did not display the sentences. Instead, it consisted of a numbered list of lines on which the listener could write a single word.

When conducted this way, to perform the test successfully it was necessary for listeners to segment the continuous stream of the sentence heard. The task was considered to be a more valid test of segmentation skills (albeit, also memory) than
the Certificate of Advanced English test used in the previous experiment, which also tested higher level processes.

Evidence for the construct validity of a test comes from content, criterion-related (Hughes, 2003) and cognitive (Field, 2011) validity. A test has high content validity if it includes a representative sample of the structures relevant to the construct being tested. The listening cloze test would provide a representative sample of the structures involved in real-life segmentation, including the phonotactic and stress shape contexts that are trained by the word-spotting method. A test has high criterion-related validity if its results correspond to other independent and highly reliable measures of the test-takers’ ability. Templeton (1977) established criterion-related validity for the ‘bleep test’ version of the listening cloze test when he found a high correlation between students’ performance on the test and teacher and peer assessments of their listening proficiency.

The aspect of cognitive validity described by Field (2011) that is relevant to the context of this study is ‘similarity of processing’, which concerns whether the test requires test-takers to apply the cognitive processes that would be used in relevant real-life listening. Successful completion of the listening cloze test requires the authentic listening processes of segmentation of the string and lexical activation. An indication of the cognitive validity of listening cloze activities generally is that Field (2008b) and Vandergrift and Goh (2012) recommends adapting the traditional cloze test (or ‘gap filling’ activity) to have students find and transcribe more than one word in a passage as a speech segmentation training exercise. However, the test does contain an additional cognitive component. Listeners need to commit a word-by-word representation of the string to memory so that the missing word may be noticed on the subsequent presentation of the sentence.
As the ‘bleep test’ form of the listening cloze test meets the requirements of content and criterion-related validity, it can be said to possess important elements of construct validity. However, the memory requirements of the test may pose a challenge to its cognitive validity.

6.2.2 Partial Dictations: Dialogues

So as not to rely on one test, a second pre- and post-test was run alongside the listening cloze tests. Dictations were chosen as they require accurate segmentation of the speech stream. Partial dictations, in which some of the aural text is provided in writing on the answer sheet, were used as they were thought to be less demanding than full dictations, and therefore to be potentially more useful in measuring training effects. Naturally, as partial dictations were being used as one of the pre- and post-tests in this experiment, dictation training could not be used as a training program.

Partial dictations have been recommended in the literature by Johannson (1973) as being economical and natural, and likewise by Oller (1979) as mirroring the requirements of natural spoken discourse. Hughes (2003) recommended giving partial dictations to test-takers with lower proficiency as they provide support in enabling the listener to keep oriented within the passage. In addition, Cai (2013) found that partial dictations were high in reliability and validity. Partial dictations were found to have internal consistency, with coefficient alphas between 0.74 and 0.81 for 10 scoring units (see also Cai, 2014). Validity was demonstrated though confirmatory factor analyses comparing partial dictations with a more established assessment technique: gap filling on summaries and constructed responses (in which listeners hear a text and must finish a sentence with a statement true of the text). The analyses showed that the two tests measure the same abilities.
Buck (2001) states that longer dictation chunks test working memory in addition to language proficiency. More advanced listeners are able to hold more words in short-term memory before grouping them together into meaningful units. The content of the chunks is remembered more readily than the precise words and is sometimes paraphrased. If the dictation segments are very short and do not challenge working memory, however, Buck suggests that the activity becomes a test of word recognition. The shorter chunks presented in partial dictations could thus meet the requirements of the pre- and post-test, with segmentation and lexical access removed from higher level processes.

The dictation test would contain a representative sample of the structures relevant to speech segmentation and to the experimental training method and the vocabulary used would be appropriate to upper intermediate English learners. The partial dictations would thus have high content validity. Criterion-related validity could not be established prior to the pilot test.

In order to perform the partial dictations, participants would need to engage in similar cognitive processes as those that apply in real-life listening: prelexical processing, normalisation and segmentation of the incoming speech stream and subsequent competition between candidate words. Written context is provided in the partial dictations, but context is also a typical component of authentic listening situations. Like the listening cloze test, there would be a memory component to the partial dictations. Yet having the chunks replayed, with the same task being expected in response to each presentation, would assist listeners recall any words they might not have recognised, in the same way that listeners could seek clarification when conversing with a real-life partner or replay an online recording (Field, 2011). For these reasons, the partial dictations have a good level of cognitive validity.
6.3 Method

6.3.1 Participants

Participants were 10 self-selected students in the highest level of the English for Academic Purposes course (at that time, EAP III) at UWSCollege. They received payment for their involvement in the course at a rate of $25 at the completion of each training session if each day of each session was completed. If participants missed a training day in any session they would receive $5 for each hour completed in that session.

The participants were of Chinese (4), Vietnamese (2), Japanese (1), Peruvian (1), Turkish (1) and Saudi Arabian (1) language backgrounds.

6.3.2 Stimuli

Two training programs were planned over a four week training session: word-spotting and connected speech processes training. The training materials for these two programs were the same as in Experiment One.

6.3.2.1 Listening Cloze Test

_Sentences_

150 non-contextual, unrelated sentences were planned in total – 50 per test. As the experiments for this project were being conducted at UWSCollege and needed to fit in with their term schedules, there was limited time between experiments to create new stimuli. Sentences recorded for use in the tests were selected from a number of sources: Akker and Cutler (2003), Cutler and Darwin (1981), the TIMIT Acoustic-Phonetic Continuous Speech Corpus (Garofolo, et al., 1993), and the Australian National Database of Spoken Language corpus (ANDOSL) (Millar, Vonwiller, Harrington & Dermody, 1994). The sentences were adapted when necessary, to reduce
their length or decrease the complexity of their vocabulary. Others containing collocations that were too predictable were not selected. See Appendix H for a list of sentences used in the pre-test, including their original sources.

Existing recordings of sentences originally used in Nazzi, Jusczyk, & Johnson (2000) were also sourced. Only recordings in which sentences were presented as clear, continuous streams of speech were selected. The sentences used were recorded by three female speakers in the general Australian English accent, including the author. Recordings were made digitally to computer using Cool Edit 2000 software and an EDIROL UA-25EX USB audio interface. A Rode NT2 condenser microphone was used.

Design of Bleep Sentences

In the pre- and post-tests, each sentence would be played once in full and then again with one word bleeped out. An electronic ‘bleep’ was created and then inserted into each sentence. The bleep sound was created using Audacity software, version 1.2.6. It took the form of a 450Hz sine wave of 320 ms duration with amplitude at midpoint of 70.97 dB and a 50 ms fade in and fade out. The bleep was saved as a .wav file.

A selected word in each sentence was cut from the stream of speech and then the bleep inserted where the word had been. The words were cut from each sentence using Praat software (Boersma, 2001), with care taken to analyse the waveform and spectrogram to find the closest approximation to the start and end of each word. Words featuring a great deal of assimilation (for instance, lenition of the /t/ in highest before the /r/ in rate in ‘highest rate’, forming the affricate [tʃ]) or overlap between

---

15 My thanks to Dr. Michael Tyler for these sentences.
segments at word boundaries were not selected. The start and end of the words were cut at the nearest zero crossings.

The two separated parts of the sentence and the bleep were then concatenated in Praat with a 0.01 s cross-fading overlap. If a word bordering another to be replaced by a bleep contained clear transitions into or out from that word, the overlap was set to 0.5 instead of 0.01 s. For instance, the in the sequence ‘the rising’ contained steeply falling F2 and F3 values due to the following /r/, and an overlap of 0.5 was also set to avoid a complete velar pinch in the vowel in to in the sequence ‘to carry’.

When function words with reduced vowels bordered the bleep, they sometimes virtually disappeared. This would not create problems for listeners as they would hear the sentence in full before hearing it again minus one word in principle.

The quality of the resulting bleep sentences was very good, with no clicks or discontinuities, and no predictability due to context.

**The Selection of Words to Replace with Bleeps**

It was tempting to have the bleep replace the least predictable word in the sentence, but over several trials this location itself would have become highly predictable. In order to achieve balance, the choice of word to be ‘bleeped’ was made through rotation through parts of speech. Only content words were selected: nouns, verbs, adjectives and adverbs.

The position of the word in each sentence was also varied in order to avoid predictability. Given the memory element to the task, primacy and recency effects might be expected, with performance better at or near each end than in the middle (Murdock, 1962). However, these processes would occur independently of training.
type and the position of the test in the course of training, and would be unlikely to
directly affect results.

The Japanese dictionary used earlier (Shibata, et. al., 1988) to place the word-spotting
material into difficulty sets was now used to assess the frequency of each word
selected to be bleeped out. If the word was not rated or was given a very low
frequency rating, it was not selected. Words were chosen if they were homophonous
with higher frequency words, but not used in a high frequency semantic context, e.g.
crown in the context ‘crown witness’. A full list of the 50 sentences used in the pilot
test, including a note of the word selected to be replaced with a bleep and its part of
speech, can be found in Appendix H.

6.3.2.2 Partial Dictations

Three partial dictations were planned; one for the pre-test and one for each post-test.
In order to facilitate generalisation, the participants were exposed to a number of
different voices in both the word-spotting and connected speech processes training.
As mentioned previously, the listening cloze pre- and post-tests featured three
recorded voices. For this reason, dialogues were used in the partial dictations.

The dialogue used in the pre-test was taken from Listening to Australia: Intermediate
(Brawn, 2002), published by the NSW Adult Migrant English Service (AMES). The
dialogue took the form of a three minute radio interview between two speakers (one
with a general Australian English accent and the other a cultivated Australian English
accent) on the topic of obesity in children (see Appendix I). Chunks of dialogue were
given in writing on the answer sheet and other, usually smaller, chunks were replaced
by blank spaces where participants could write what they had heard. The potentially
more difficult sections, due to low-frequency vocabulary or being said more quickly,
and hesitations and false starts were given in writing.
6.3.3 Procedure

6.3.3.1 Pre-test

In Experiment One, the pre- and post-tests were conducted in a room with a high cathedral ceiling, which caused sound reflection. The pilot test of the new pre-test took place in a room with a low, flat ceiling and acoustic quality was improved.

The participants first heard the listening cloze test. The first four sentences were used as practice trials and the author checked for understanding of the task requirements before proceeding to the first test item. Individual sentences were initially played in their entirety and then repeated with one word replaced by a bleep. Participants were given time to write single word answers on their answer sheets. The answer sheet consisted of two pages of numbered lines, on which participants would be asked to write one word per line.

They were then given the partial dictation. The dialogue was played once in full and then again with pauses after each section which corresponded to a blank space on the answer sheet (or a number of times within these sections if they were long). Participants were given the chance to write down what they had heard. The whole passage was then played again so that the participants could check what they had written and make any necessary changes.

Both pre-tests were played using iTunes software on a Lenovo ThinkVantage T61p laptop, through Panasonic SB-PM21 speakers.

Participants reported that they found the listening cloze test very hard, but had greater confidence in their performance of the partial dictation task.
6.3.3.2 Training Programs

The 10 participants were divided into two groups. Five commenced training in the word-spotting method and five in connected speech processes.

**Word-Spotting**

Due to the restrictions placed by UWSCollege on external software, the program needed to be installed at the start of each day's training. A trial of the software in the days preceding the commencement of the pilot test had been successful, with three computers installing DMDX and five computers running it simultaneously. However, when the five participants in the training program tried to install DMDX simultaneously, it slowed down and then froze. Once it had frozen the computer was unusable and the participants couldn't log off to log in on another computer. The next day, the installation of the program was staggered - one student at a time. However, this was too time-consuming to be considered successful.

The problems continued, despite the efforts of the college's technical officer to solve them, over the following days. For this reason, the word-spotting training could again not be run on the computers. Instead, the training was run on a Lenovo ThinkVantage T61p laptop, through Panasonic SB-PM21 speakers, as it was in Experiment One.

**Connected Speech Processes**

The connected speech processes training was run as per Experiment One, again on a Lenovo ThinkVantage T61p laptop, through Panasonic SB-PM21 speakers. As it was not dependent on the college computers, the training sessions ran very smoothly.
6.4 Cessation of Pilot Test

The pilot test was brought to an end at the completion of the first week of training. The reasons for the decision to truncate the planned experiment were that there was already good cause to believe the listening cloze pre-test had been too difficult and a trial of the word-spotting training had revealed that the technical problems at UWSCollege had not yet been resolved. Ending the pilot test at this point allowed more time to remedy the problems in the hope of successful resumption in the next teaching session at UWSCollege.

6.5 Results

6.5.1 Listening Cloze Test

An analysis of the listening cloze pre-test confirmed that participants had found it prohibitively difficult. Of the 46 test items, 20 (43%) received incorrect or blank answers from all participants. The highest correct score was 35%, achieved by two participants; the lowest was 7% (3 out of 46 words correct). The mean was 18% ($SD = 9.59$).

6.5.2 Partial Dictations

6.5.2.1 Notes on the Marking Scheme

In line with modern conventions in dictation marking (Heath, 1985; Oller, 1979), the total number of words in the text was counted and one mark allowed for each (203 words in total). Spelling errors were not marked as incorrect where the word written was a clearly recognisable attempt at the word dictated, for example carolis for ‘calories’.
Some deviations from general purpose dictation marking conventions were made in an effort to achieve validity in scoring, an important component of construct validity (Hughes, 2003). In doing so, it could be ensured that both the dictation task and the scoring system tested speech segmentation. Particular errors were allowed, whereas others were disallowed and ignored as incorrect responses or, where appropriate, marked as missegmentations.

Spurious words were not counted or subtracted from the total score. Words are often inserted when dictation takers attempt to reanalyse what they think they have heard and make it conform to their ideas of correct grammar. These words are often articles. The exception was when such words were clearly inserted due to missegmentation of the stream, such as sector is in place of ‘sectors’.

As mentioned previously, listeners remember the content of longer chunks of dictation rather than individual words and tend to paraphrase (Buck 2001). For this reason, kids was accepted in place of ‘children’ and the highest level of obesity as ‘the highest rates of obesity’.

Highly common vowel errors that did not affect segmentation were allowed; for instance, we seat them down was allowed for ‘we sit them down’; so too circle for ‘cycle’. An instance where an error was made with one word medial consonant (centres for ‘sectors’) was marked as correct. Cases such as these, suggesting problems with the perception of word medial segments, are irrelevant to the assessment of segmentation ability.

Regular plural, past tense and adjectival /d/-ending errors were allowed in most cases, for example, ‘concerned about’ written as concern about. The presence of the correct word in these cases shows that the word was heard and obviously segmented from those around it. Errors such as these are more likely a reflection of the participant’s
level of grammatical proficiency. This study is not concerned with the participants’ ability to analyse each sentence as a linguistic unit (although dictation can be used in this way).

A caveat to this point is when the plural or tense markers are given to the following word (for example, ‘computers for’ written as computer so, and ‘changed our’ written as change their) or shared with it (‘heart soap’ written as hearts soap). Cases such as these are examples of missegmentation.

The absence of the plural /əz/ or past tense /əd/ was also classed as a missegmentation if it was followed by the insertion of a word; for example, change is instead of ‘changes’ and disease has in place of ‘diseases’.

6.5.2.2 Partial Dictation Results

The participants completed the partial dictation task far more successfully than they did the listening cloze test. The mean was 45.3% (s = 14.31). Figure 6.1 shows the frequency distribution of the scores.

![Figure 6.1 The frequency distribution of the partial dictation scores. N=10](image-url)
6.5.2.3 Missegmentations

Each participant produced missegmentations in the partial dictation task. The number of missegmentations ranged from 3 to 10, with an average of 6.2.

Correlations

Interestingly, there was a significant negative correlation between the performance of the participants in the listening cloze test and the number of missegmentations they made in the partial dictation task, \( r(8) = -0.672, p = .03 \). The higher the percentage of correct words in the listening cloze test, the lower the number of missegmentations in the partial dictation task. A negative correlation between the percentage of correct words in the partial dictation and the number of missegmentations a participant made was not significant, \( r(8) = -0.517, p = .13 \).

Clearly, the two scores were not independent, with missegmentation being one of many factors which may impact upon dictation scores. The objective was to establish just how strong an impact missegmentation may have.

6.5.3 Correlations between Listening Cloze and Dictations

There was a strong correlation between the scores on the listening cloze test and the partial dictations, \( r(8) = .819, p = .004 \), showing that the better a participant scored on the listening cloze test, the better they performed at the partial dictations.

6.6 Discussion

6.6.1 Listening Cloze Test

The negative correlation between the number of correct words in the listening cloze test and the number of missegmentations in the partial dictation indicate that accurate segmentation ability is required in order to complete the listening cloze test.
The participants who were able to recognise more ‘bleeped out’ words in the listening cloze test made fewer missegmentation errors in the partial dictation. It is encouraging to think that the rationale for using the listening cloze test had been sound.

However, the listening cloze test was deemed inappropriate for further development and use. It was excessively difficult for the participants – recall that 43% of the words were not recognised correctly by any of the participants - and to use it in subsequent experiments would be to engender the risk of participants performing ‘at floor’ in each post-test.

The listening cloze test was subsequently given to a native English listener, who achieved 100% success in the task. It is important to understand why the test might have been so easy for a native listener but so difficult for English learners. Some aspects that may have contributed to the difficulty of the task for second language listeners are the length of the sentences, the speed of the speech, the non-contextual nature of the task and sentences themselves, and the total cognitive load on the listener. The longer sentences of the task may have proved a memory challenge. The participants may have actually segmented them quite well, but had trouble retaining the words in their short term memory as they segmented the rest of the sentence. Recall that it was the extraneous demands on memory that presented a challenge to the test’s cognitive validity.

Upon more reflective analysis of the recordings, it was noted that some sentences were said quite fast; perhaps faster than one would normally speak. The longer the sentence was, the faster it was said. This is a phenomenon the author has noted in numerous recordings of numerous speakers. When being recorded, speakers
establish a duration pattern and make all items conform to it as they move through the list before them. This was an issue to be aware of and avoid in later recordings.

The lack of context between sentences would have created a challenge for the participants, but the ‘strangeness’ within each sentence would have made them doubly hard to segment. A sentence like, “The rising price of boxes worried the dog food manufacturer” is unexpected - which was the point - but the bizarreness of the sentence may have created problems for second language listeners. It is likely that they were still trying to establish whether they had heard the start of the sentence correctly while the rest of the sentence was demanding their attention. Perhaps the fact that there are few contextual cues in these sentences, while not an issue for speech segmentation and lexical activation, does impact negatively on working memory.

The native listener who sat the test said it took a lot of effort and concentration to perform the task, so for a second language learner it must have been very tiring. The above considerations, taken together, suggest that the cognitive load may have been too great. It was decided that future materials should feature shorter, slower sentences. However, the contextual cues would not be increased, in order to avoid priming or guessing of words.

6.6.2 Partial Dictation

The outcome of the partial dictation test was more promising. The percentage of correct words given by participants resulted in a bell shaped curve, with a mean score nearing 50%. This pre-test result gave promise that there would be room for improvement in the task when used as a post-test. In addition, the strong relationship between performance on the partial dictations and the listening cloze test provided support for the task having criterion-related validity, and as such,
construct validity. The relationship between the number of missegmentations in the partial dictation and the number of correct responses in the listening cloze test was encouraging. It provided support to the idea that tallying missegmentations in dictations in subsequent experiments would be a valid endeavour which could measure changes in segmentation ability.

6.7 Future Directions

Dictation would be used as the pre- and post-tests in subsequent experiments, with a change to the task format. Instead of dialogues, the shorter listening cloze test sentences (without the bleeps) would be used as dictations and others newly written and recorded. The potential to measure improvement in segmentation ability in the form of higher dictation scores and lower missegmentation counts over the course of the tests was recognised. Any correlation between the number of missegmentations made in the dictations and participants’ performance in the word-spotting task would also be studied in future experiments.
Chapter 7

Experiment Three

7.1 Aim, Design and Hypotheses

As in previous experiments, the aim of Experiment Three was to assess the efficacy of the word-spotting task as a method of training second language speech segmentation. The study followed a pre-test - treatment - post-test design, with training in connected speech processes used as a control. Experiment Three included an improved series of post-tests. In addition to dictation post-tests, which took place at the completion of each ten day training program, participants were given ‘mini-dictations’ at the end of each day’s training session in order to track their learning throughout the training. The hypotheses were that:

1. Participants who received connected speech processes training would show improvement in dictation scores and progressively fewer missegmentations, in keeping with empirical evidence of its efficacy (Brown & Hilferty, 1986; Kennedy & Blanchet, 2013).

2. As the word-spotting training has been designed to be segmentation-specific, the group receiving this training would also show improvement in dictation and missegmentation scores.
3. Implicit training in word-spotting may lead to greater improvement than the explicit, metalinguistic training in connected speech processes.

7.2 Method

7.2.1 Participants

Participants were 14 (8 female and 6 male) students in the highest level English for Academic Purposes class at UWSCollege. To gain entry to this course, students must have an IELTS score of 6 or more, with a minimum score of 5 in all test areas, or equivalent score in another recognised test. The participants’ first languages can be grouped as Korean (5), Chinese (4; Mandarin, Southwestern Mandarin, Shanghainese and Hokkien dialects, respectively), Indo-Aryan (2; Punjabi and Bengali, respectively), Vietnamese (2) and Tagalog (1).

Participants were randomly allocated to training groups. The composition of the groups is presented in Table 7.1.

<table>
<thead>
<tr>
<th>Group</th>
<th>Number</th>
<th>Language</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1C2</td>
<td>7</td>
<td>Chinese – 3</td>
<td>Female – 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Korean – 3</td>
<td>Male – 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indo-Aryan - 1</td>
<td></td>
</tr>
<tr>
<td>C1W2</td>
<td>7</td>
<td>Chinese - 1</td>
<td>Female – 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Korean - 2</td>
<td>Male – 5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Indo-Aryan - 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Vietnamese – 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tagalog – 1</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1 The composition of the two training groups. The acronyms refer to the order in which the groups received the training.

The study was advertised to the students during class time by the author. Those interested later attended an information session run by the author in which the
details of the training were explained. Students wishing to participate wrote their email addresses on a list and were then contacted by the author and given their training group and time details.

Participants were paid $80 if they attended the pre-test, all twenty training sessions and both post-tests. If they missed any training sessions, they were paid a rate of $2.50 per session, instead. Table 7.2 displays the completion rates per group per session. Two additional students completed the pre-test, but did not continue with the training.

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th>Session 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WS</td>
<td>CSP</td>
</tr>
<tr>
<td></td>
<td>Group W1C2</td>
<td>Group C1W2</td>
</tr>
<tr>
<td>100%</td>
<td>2</td>
<td>100%</td>
</tr>
<tr>
<td>75 - 99%</td>
<td>2</td>
<td>75 - 99%</td>
</tr>
<tr>
<td>50 - 74%</td>
<td>1</td>
<td>50 - 74%</td>
</tr>
<tr>
<td>0 - 49%</td>
<td>2</td>
<td>0 - 49%</td>
</tr>
<tr>
<td></td>
<td>Group C1W2</td>
<td>Group W1C2</td>
</tr>
<tr>
<td>100%</td>
<td>0</td>
<td>100%</td>
</tr>
<tr>
<td>75 - 99%</td>
<td>2</td>
<td>75 - 99%</td>
</tr>
<tr>
<td>50 - 74%</td>
<td>4</td>
<td>50 - 74%</td>
</tr>
<tr>
<td>0 - 49%</td>
<td>1</td>
<td>0 - 49%</td>
</tr>
</tbody>
</table>

Table 7.2 The percentage of each training session completed by the number of participants in each training group.
7.2.2 Stimuli

7.2.2.1 Training Programs

The training programs run in Experiment Three were the word-spotting and connected speech processes training. As dictation tasks were used as the pre- and post-tests in this experiment, dictation training was not offered.

The materials used in Experiment Three were the same as those used in Experiment One. However, the method of delivery of the word-spotting training was changed. In this experiment, participants were able to run the training on the computers in a self-study format as it was initially designed to be delivered.

7.2.2.2 Pre- and Post-tests

Dictations

Following the successful use of partial dictation in Experiment Two, a new format of pre- and post-tests using dictation was designed for this experiment. The tests consisted of 40 unrelated dictation sentences, which can be found in Appendix J. These sentences included those used in the listening cloze test in Experiment Two (re-recorded if necessary), further sentences selected from Cutler (1976) and Cutler and Fodor (1979), and newly written sentences, some adapted from feature articles on the Australian Broadcasting Commission’s website: www.abc.net.au.

Buck (2001) recommends offering a variety of difficulty levels in dictations by manipulating the length of ‘chunks’ to be transcribed. Shorter dictation chunks should appear at the beginning of tests, with chunks getting progressively longer.
Accordingly, there were three components to the new test: a practice phase of ten sentences, followed by 15 ‘partial’ dictation sentences and a further 15 ‘whole’ dictation sentences. The practice phase comprised:

- Two sentences that were written in full on the answer sheet. The participants were asked to simply listen to these sentences.
- Three sentences in which one word was missing.
- Five sentences in which two words were missing.

In Appendix J, the sentences are listed in the order they appeared in the pre- and post-tests and the words that were missing on the answer sheet are underlined.

The purpose of the practice phase was to allow the participants to ‘tune in’ to the sentences. They were exposed to the speaker’s voice, the rate of speech, the lack of context and the level of vocabulary difficulty involved. They were also given practice at performing the task with gradually increasing complexity. These sentences were not considered tests of segmentation ability as the segmentation had already been done for the listener.

In the partial dictations set, the beginning of the sentence was given and the following words left blank. These sentences provided some context. The next 15 whole dictation sentences were left completely blank on the answer sheet and listeners were asked to transcribe the entire sentence.

**Mini Dictations**

In addition to the pre- and post-test dictations, participants were given ‘mini-dictations’ – ten sentences of which five were partial dictations and five were whole dictations – at the end of each training session. The mini-dictations were designed to measure gradual progress in segmentation skills throughout the training,
compensating for the fact that some participants might attend the training, but not the post-tests (as had occurred in Experiment One). See Appendix K for a full list of mini-dictations used in Experiment Three.

122 sentences were created for the dictation post-tests and mini-dictations. They only contained words that would be expected to be known by students at the upper intermediate proficiency level (as ascertained in reference to Shibata et al., 1988).

7.2.3 Equipment

The audio files for the pre-test, post-tests, mini-dictations and connected speech processes training were played through Windows Media Player 9 series software on a Lenovo ThinkVantage T61p laptop and through Panasonic SB-PM21 speakers. The room used had a flat ceiling and there was no problem with sound reflection.

The word-spotting training was conducted on custom computers supplied by a Sydney company, Impact Systems Technology Pty Ltd. The processor model was Intel i5, with processor speed of 3.2ghz and 4gb of RAM and the operating system was Windows Vista. 1 BenQ E2220HD 22” LCD monitors were used. The training was displayed and responses collected by DMDX software (Forster & Forster, 2003). Participants wore Logitech H110 Stereo Headsets.

The audio files for both the word-spotting training and the dictations were normalised using Audacity software, version 1.2.6.

7.2.4 Procedure

A total of 14 participants commenced training on the first day – seven in each group. The two groups received both of the training treatments, but in different orders. The groups will be referred to here as W1C2 and C1W2, according to the order in which they received two training methods. Group W1C2 received word-spotting training in
session one and connected speech processes training in session two, while group C1W2 received the training in reverse order.

Each training program took place over three weeks from Monday to Thursday. An information session took place on the first day. On week three, the Thursday session was the post-test. This gave nine days of 30 minute training sessions before each post-test.

As in Experiment One, training took place at the end of the day’s courses at UWSCollege, starting at 2:10 in the afternoon. The training programs were run by the author, consecutively.

The word-spotting training was run on the computers in a self-study format. This often extended the amount of time needed to run the training. Due to server issues at UWSCollege, each participant needed to install the program each day. In addition, all participants had to be finished the training before the mini-dictation task could be conducted. As the word-spotting training ran to an unpredictable schedule, it was held after the connected speech processes training each day instead of alternating with it as had occurred in Experiment One. The training schedule is shown in Table 7.3.

<table>
<thead>
<tr>
<th>Group W1C2</th>
<th>Group C1W2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Week 1</td>
<td>Dictation Pre-Test</td>
</tr>
<tr>
<td>Weeks 2 - 4</td>
<td>Word-Spotting</td>
</tr>
<tr>
<td>Week 4</td>
<td>Dictation Post-Test 1</td>
</tr>
<tr>
<td>Week 5</td>
<td>No Training</td>
</tr>
<tr>
<td>Weeks 6 - 8</td>
<td>CSP</td>
</tr>
<tr>
<td>Week 8</td>
<td>Dictation Post-Test 2</td>
</tr>
</tbody>
</table>

*Table 7.3. Training schedule for Experiment Three*
7.2.4.1 Pre-Test

Participants in each group received the pre-test simultaneously. Each section of the pre-test was explained and the author answered all requests for clarification prior to the test. Each training and test sentence was played twice and the participants given time to write their answers.

7.2.4.2 Training Programs

Word-spotting

On the first day of the word-spotting training, participants were given a rationale for the training and were shown how to install and use the DMDX software.

Participants sat at individual computers wearing headphones. The presentation of each stimulus item proceeded in the following way. A first screen appeared displaying the words, “GET READY”. An audio file of a nonsense string was then played, for example, “nuhten”. Participants then typed the target word they heard (the correct answer, in this case, being ‘ten’) and pressed ‘enter’. A feedback screen was then shown, with the words, “You heard nuh-TEN” with the target word in capitals and the nonsense syllable in lower case. Instances of schwa were written as ‘uh’ to avoid confusion with whole vowels, and thus sometimes real English words (in this case, ‘net’). When participants were ready to proceed to the next item, they pressed the space bar and the words “GET READY” again appeared. Participants continued on with the task for the 100 items in each day’s training session.

When each participant had completed the training, the words, “THE END” appeared on their screen. The author then copied the output zil file to USB and saved it under the participant’s ID code and the day’s training set.

At the end of each day, the author tallied the number of correct responses for each participant. The participants were given their individual results on a slip of paper at
the beginning of the next class. The results were confidential. However, the highest score and the class average were revealed to the group each day in order to motivate the participants to improve.

*Connected Speech Processes*

The training in connected speech processes proceeded in the same way as it had in Experiment One. Participants were given feedback on the accuracy of their responses in order to motivate improvement.

### 7.2.4.3 Post-test Instruments

*Daily Mini-Dictations*

When each group had completed a day’s training session, they were given a mini-dictation test. Each sentence was played twice and participants were given time to complete it before the next was played. Feedback was not given on the correct sentences or the participants’ performance.

*Post-tests*

The post-tests took place the day after the final training day in each program. They followed the same procedure as the pre-test and both groups of participants received each simultaneously. Participants were paid according to their attendance rate at the end of each post-test. The questionnaire was given at the end of post-test 2.

### 7.3 Results

Unless otherwise specified all data was analysed using a repeated measures analysis of variance (ANOVA), with training group as the between-groups factor, and an alpha level of .05. Assumptions of normality and homogeneity were satisfactory unless otherwise discussed.
Only the statistics for significant effects and interactions will be presented here. The complete statistical data can be found in Appendix L.

7.3.1 Dictations

7.3.1.1 Pre-test

An independent t test was conducted on the pre-test dictation results to assess potential differences between the groups prior to training. While the mean scores of the W1C2 group were higher than those of the C1W2 group, the t test indicated that the difference between the two groups was not statistically significant in either the partial dictation component of the pre-test (W1C2: \(M = 67.60, SD = 9.53\); C1W2: \(M = 58.00, SD = 15.47\)), or in the whole dictations (W1C2: \(M = 53.20, SD = 8.84\); C1W2: \(M = 49.43, SD = 12.42\)). The results suggest, however, that participants found the partial dictation component easier to complete than the whole dictation component.

Missegmentations

Independent t tests were conducted on the number of missegmentations in the pre-test. Again, there was no significant difference between the groups in the partial dictations, (W1C2: \(M = 4.80, SD = 1.30\); C1W2: \(M = 4.57, SD = 1.27\)), nor the whole dictations (W1C2: \(M = 9.60, SD = 2.19\); C1W2: \(M = 8.00, SD = 3.00\)).

Correlations

There was no correlation between combined group performance in the dictation and missegmentation scores for the partial or whole dictations.

7.3.1.2 Pre-test – Post-test 1

The first set of analyses was conducted on the results of those who completed 80% or more of the training – four participants in each group. Group W1C2 contained two males and two females. The first languages were Korean (2), Tagalog (1) and
Vietnamese (1). Group C1W2 contained four females. Their language backgrounds were Korean (2), Punjabi (1) and Mandarin (1).

![Figure 7.1](image.png)

*Figure 7.1 Mean correct scores in the partial and whole dictation components from the pre-test to post-test 1. In the whole component, the scores of the group which received word-spotting training improved, while the scores for those who received training in connected speech processes declined. N=8*

As can be seen in Figure 7.1, the partial dictation scores of each group decreased slightly from the pre-test to post-test 1. However, the scores were very similar across testing times. The slight decrease may have been caused by a flaw in the design of the post-tests in that the number of dictated words was not held constant. While the partial dictations in the pre-test averaged 5.5 words, in post-test 1 they averaged 8 words. The whole dictations in the pre-test averaged 10 words; in post-test 1 they averaged 12 words.

The results of the analysis of the whole dictation component show a different pattern. Despite post-test 1 being slightly more difficult than the pre-test, Group W1C2 (which received training in the word-spotting method) improved on their pre-test scores, as Figure 7.1 shows, while the scores of Group C1W2 (which received training in connected speech processes) declined.
There were no significant effects or interactions in the partial or whole dictation components for the pre-test to post-test 1.

**Missegmentations**

Both groups made more missegmentations in post-test 1 than they had in the pre-test. (Partial dictations: W1C2 +0.75, C1W2 +4.00; whole dictations: W1C2 +0.5, C1W2 +0.75.) However, post-test 1 contained 54 more words than the pre-test, providing more chances for missegmentation to occur.

In order to provide a more accurate analysis of the missegmentation data, the mean number of missegmentations made was expressed as a percentage of the total number of words. As Figure 7.2 shows, Group W1C2, which received word-spotting training, made fewer missegmentations in the partial dictation component of post-test 1 than they had in the pre-test, while the group that received training in connected speech processes made more missegmentations. In the whole dictation component, the percentage of words missegmented by both groups decreased very slightly and the data shows a small advantage for Group W1C2.

There were no significant main effects or interactions for the partial or whole dictation components.
Correlations

Pearson product-moment correlations were performed between the post-test 1 dictation and missegmentation scores. A strong, negative relationship between partial dictation scores and partial dictation missegmentation scores was significant, $r(8) = -0.83$, $p < .01$, showing the higher the number of missegmented words, the lower the overall dictation score. There was no correlation between the percentage of correct words in the whole dictations and the percentage of missegmentation errors in the whole dictations.

7.3.1.3 Pre-test – Post-test 2

Only five participants in total completed more than 80% of the training in session two, so the analyses are based on those who had completed 60% or more of the training – four participants in each group. Group WiC2 consisted of two males and two females. Their language backgrounds were Korean (2), Mandarin (1) and
Vietnamese (i). Group C1W2 consisted of four females. Their first languages were Mandarin (2), Korean (1) and Punjabi (1). The mean scores across the pre-test, post-test 1 and post-test 2 cannot be directly compared as the analyses were composed of the results of different participants.

The mean percentages of correct answers in the partial and whole dictation components of each test appear in Figure 7.3. In the partial dictations, the performance of both training groups declined. There was a significant effect of ‘test number’, $F(1, 7) = 10.14, p = .02$. For the whole dictation component the results are very different, with the mean scores of both training groups increasing in post-test 2. The main effect of ‘test number’ was again significant, $F(1, 7) = 6.16, p = .048$. There was a significant main effect of ‘component’, $F(1, 7) = 11.72, p = .014$, and interaction between ‘test number’ and ‘component’, $F(1, 7) = 27.81, p = .002$.

There was no significant effect of ‘training group’, nor interaction of ‘test number’ and ‘training group’ for the partial or whole dictation components.

*Figure 7.3* Mean percentage of correct words in the partial and whole dictation components of the pre-test and post-test 2. In the partial component, there was a significant (at the 0.05 level) decline in the scores of both groups. In the whole component, results significantly improved for both groups (at the 0.05 level). $N=8$
Misssegmentations

Due to the differences in test length across the components of the pre-test and post-test 2, the misssegmentations were analysed as a percentage of the total number of words in each component. The mean number of misssegmentations made by each training group in the partial dictation component rose from the pre-test to post-test 2 and there was a significant main effect of ‘test number’, $F(1, 7) = 6.67, p = .04$. The main effect of ‘training group’ and the ‘training group’ by ‘test number’ interaction were not significant. The mean scores for each group can be seen in Figure 7.4.

![Figure 7.4](image)

Figure 7.4 The mean number of misssegmentations made in the pre-test and post-test 2. Again, the lower the number, the better the performance. In the whole component, the number of misssegmentations made by W1C2, which received word-spotting in session 1 and connected speech processes in session 2, decreased, while C1W2’s misssegmentations increased. N=8

The mean number of misssegmentations made in the whole dictation component of the pre-test and post-test 2 was also analysed. As can be seen in Figure 7.4, Group W1C2, which had received word-spotting training in the first session and training in connected speech processes in session 2, made fewer misssegmentations in post-test 2. Whereas Group C1W2, which had received connected speech training in the first
session and word-spotting training in session 2, made more missegmentations. This particular configuration of Group W1C2 had made substantially more missegmentations than Group C1W2 in the pre-test, so the difference in performance is noteworthy. However, there were no significant main effects or interactions.

**Correlations**

Pearson product-moment correlations were performed between the post-test 2 dictation and missegmentation scores. There was no significant correlation between partial dictation and partial missegmentation scores, nor whole dictation and whole missegmentation scores.

### 7.3.2 Mini-Dictations

#### 7.3.2.1 Session 1

An analysis was conducted on the results gained from the daily mini-dictations for those participants who attended 80% or more of the training (four participants in each group). As periods of consecutive days of attendance occurred inconsistently within and between groups, the data was analysed as a function of the number of training days attended by each participant rather than the training day itself. For instance, one participant missed day two of the training, so her result on day three was analysed as ‘test number two’ and so on.

The main effect for ‘test number’ was significant, $F(5, 7) = 3.24$, $p = .02$. Pairwise comparisons showed a significant effect for ‘test number’ between ‘Test number 1’ and ‘Test number 4’, $p = 0.02$, and ‘Test number 1 and Test number 5’, $p = 0.05$, with the mean percentage correct being higher in Test 1 and dropping in the later tests.

Figure 7.5 shows the mean mini-dictation results for each group over six testing times. Each group’s scores drops after day 1. The mean scores for the group that received
word-spotting training are consistently higher than those for the group that received training in connected speech processes. However the main effect of ‘training group’ and the ‘training group by test number’ interaction were not significant.

Figure 7.5 Mean mini-dictation results for session 1 of the training. The mean scores for the group that received word-spotting training start to rise from Time 5. N=8

**Missegmentations**

An analysis was conducted to compare the mean number of missegmentations occurring over the course of the first six mini-dictations across the two training groups. Again, only those four participants in each group who had completed 80% or more of the training were assessed.

As Figure 7.6 shows, after an initial rise in missegmentation scores, the group that received word-spotting training made continually fewer missegmentations over the course of the training compared to the more inconsistent performance in the connected speech processes group. However, there were no significant effects or interactions.
Figure 7.6 Mean missegmentations per group over the course of the first six tests completed by each participant. Group W1C2, which received word-spotting training in session 1, made consistent improvement from Day 2 on. N=8

Correlations

Pearson product-moment correlations were performed across the mini-dictation and missegmentation scores. In each case there was a negative correlation, but none reached statistical significance.

7.3.2.2 Session 2

The session two mini-dictations were roughly the same length as those in session one. Six sentences contained an average of nine words each and the other four contained an average of ten.

An analysis was performed on the mini-dictation scores of those participants who had completed 60% or more of the training in session two of the experiment – five participants in each group. (This mini-dictation analysis had a higher number of participants in each group than the post-test 2 analysis as not all the participants who attended the training sat the post-test.) In the second training session, the group that had received word-spotting training previously now received training in connected
speech processes and vice versa. The main effect of test number was significant, $F(5, 9) = 3.47, p = .01$. Neither the main effect of ‘training group’, nor the interaction between ‘training group’ and ‘test number’ were significant.

**Figure 7.7** The mean percentage of correct scores in each group across six testing times in session 2. The scores of Group W1C2 are again consistently higher than those of Group C1W2, but Group C1W2 made greater improvement throughout the session. N=10

Figure 7.7 displays the mean scores across the six testing times for each group. The scores rise and fall for each group, but greater improvement in mean scores can be seen for the group that received word-spotting training in this session.

**Missegmentations**

The mean number of missegmentations in the mini-dictations made by the five participants who completed 60% or more of the training in each training groups was analysed.
Figure 7.8 shows that despite the rise and fall of the scores, both groups can be seen to improve gradually over time. However, there were no significant main effects or interactions for this analysis.

**Correlations**

Pearson product-moment correlations showed a very strong negative correlation, $r(10) = -.82$, $p = .004$, between mini-dictation and missegmentation scores in the case of Test 3. No other analyses reached significance.

### 7.3.3 Word-Spotting Scores

An analysis was performed on the daily results achieved in the word-spotting training task across six training days. The two groups were combined for this analysis, creating a total number of nine participants in one group. The main effect of test number was significant, $F(5, 8) = 9.89$, $p < .001$.

Mean scores across the six testing times are shown in Figure 7.9. The scores show that performance of the word-spotting task improves with repeated exposure.
7.3.3.1 Correlations

Pearson product-moment correlations were performed on each set of data collected each training day: the score achieved on the word-spotting task and the number of missegmentations made in that day’s mini-dictation. Again, the data reflects not consecutive days, but consecutive training and testing sessions for each participant.

The correlation analysis of Day 1 showed a strong, negative relationship between word-spotting and missegmentation scores, which was significant, \( r(9) = -.84, p = .004 \) (the higher the word-spotting score, the lower the number of missegmentations made in the mini-dictations). There were no other significant correlations.

7.3.4 Questionnaire

The questionnaire given to participants in Experiment Three can be found in Appendix M. Connected speech processes training was called ‘speech awareness training’ during the experiment and in the questionnaire. That label, while not a technical term, could be more readily understood by the participants.
Table 7.4 shows the number of participants who rated each training type most helpful and least helpful.

<table>
<thead>
<tr>
<th></th>
<th>Most Helpful</th>
<th>Least Helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>CSP</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>No response</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

*Table 7.4 Questionnaire results. The number of questionnaire respondents who rated each training type as most and least helpful*

The one participant who rated the word-spotting training as most helpful stated that the training focused attention on listening to a speaker’s pronunciation. The participants who rated the training in connected speech processes as most helpful commented on the merits of studying speech as it is genuinely spoken in the Australian English accent and also mentioned its side benefit in training their own pronunciation. One such participant noted the connected speech training enhanced their speaking skills, but the word-spotting training was helpful for their listening.

The word-spotting training was criticized by some participants as not presenting real speech. As such, some participants believed that it did not provide useful listening practice. One participant said the training could be improved by giving “a whole article or short paragraph to practice”.

The participants thought the connected speech processes training program was too short and could be improved by offering more weeks of training.
The descriptors chosen by the participants for each training program are presented in Figure 7.10. The word-spotting training received more negative descriptors than the connected speech processes training.

Participants stated a clear preference for the training in connected speech processes, which was seen to be more useful. Yet, the participants were very enthusiastic and eager while undertaking the word-spotting training. For instance, some gave 'fist pumps' when they got words right, reported that it was fun and that they wanted to improve.

An important factor in the dispreference for the word-spotting training was the criticism that the training did not provide practice in listening to 'real speech'. It
seems, despite being given an information session on the first day of the training session, some participants did not understand that the goal of the training was to present repeated exposure to and practice segmenting the phonetic sequences and stress shapes present in English word boundaries. This and the request made by two participants for more detailed feedback after each training session, was informative for the planning of future training and must be considered should the task ever be developed commercially.

7.4 Discussion

Although there were no statistically significant results across the pre- and post-tests, there are positive signs that the word-spotting training led to benefits of practical importance. An important finding in this study was that the combined data arising from the word-spotting task itself indicates that repeated exposure to the task improved performance of it, with significantly higher scores being recorded for each day of training. The significant negative correlation between word-spotting and missegmentation scores on Day 1, suggests the existence of a relationship between performance of the task and segmentation skills more generally.

Moreover, the general improvement across training sessions combined with a lack of significant difference between the groups can be seen, in itself, as promising. As previously discussed, Brown and Hilferty (1986) and Kennedy and Blanchet (2013) have provided empirical evidence that training in connected speech processes improves listening comprehension and segmentation. As word-spotting training and training in connected speech processes both resulted in statistically significant improvements, it is clear that the novel method is at least as effective as the established method.
Unfortunately, with just four participants in each training group in each trial, participant numbers were too low in this experiment to make any definitive conclusions about whether either training program had a positive effect on dictation performance. Variations in the number of words across the dictation pre- and post-tests are also thought to have had an impact on results.

Far higher student numbers were expected in the following term of study at UWSCollege with the annual peak in enrolments. Recall that sample size was not a problem experienced during the validation of the task (Chapter 4) and Experiment One (Chapter 5), which took place during terms of higher College enrolments. A further trial of the word-spotting materials was scheduled for the following term, with improvements made to consistency between pre- and post-test complexity.
Chapter 8

Experiment Four

8.1 Aim, Design and Hypotheses

The aims of Experiment Four were: 1) to test the efficacy of the training programs with higher participant numbers and improved consistency between the pre- and post-tests, and 2) to seek further confirmation that the word-spotting task trains speech segmentation specifically. In order to do the latter, both groups of participants were given an additional post-test which measured reaction times in noticing newly taught words in continuous speech.

Speech segmentation has been tested in a similar fashion in infant studies by Jusczyk, Houston and Newsome (1999). They familiarised 7.5 and 10.5 month old infants with disyllabic target words and subsequently measured the length of time the infants attended to these words in passages of continuous speech. In Experiment Four of the present study, upper intermediate English learners were taught novel, low frequency words. They were then asked to press a button when they heard the words in recorded sentences and their reaction times were measured.

We hypothesised that participants who had received training in the word-spotting method would respond faster to hearing the newly taught words in continuous speech
than those who had received training in connected speech processes. Again, the hypotheses were:

1. The training in connected speech processes would lead to improved post-test and daily mini-dictations scores, in line with findings reported for English in Brown & Hilferty (1986) and Québec French by Kennedy and Blanchett (2013);

2. The word-spotting training, which is segmentation-specific, would also lead to improvement in dictation scores;

3. The segmentation-specific practice in the word-spotting training would lead to greater reduction in missegmentations than the metalinguistic knowledge imparted in training in connected speech processes.

8.2 Method

8.2.1 Participants

Participants were 31 students (23 female and 8 male) in the highest level of the English for Academic Purposes class at UWSCollege. The students who had participated in Experiment Three had graduated from the College and did not join Experiment Four.

The first languages of the participants were grouped as Chinese (21), Farsi (2), Spanish (2), Arabic (1), Bengali (1), Hindi (1), Japanese (1), Korean (1), and Cebuano, a dialect of the Visayan language of the Philippines (1). The participants grouped here as ‘Chinese’ speakers spoke Mandarin (15), Cantonese (3), Hangzhounese (1), Shanghainese (1) and Wu (1) dialects. The two Spanish speakers spoke Mexican Spanish and Peruvian Spanish dialects, respectively.

Participants were randomly assigned to the two training groups, WiC2 (n = 15) and CiW2 (n = 16), the composition of which appears in Table 8.1.
Table 8.1 The composition of the two training groups.

The study was advertised and information sessions run in the same way as in Experiment Three. The participants were also paid at the same rate and with the same conditions as in Experiment Three. Table 8.2 displays the completion rates per group, per training session.

<table>
<thead>
<tr>
<th>Group Number</th>
<th>Language</th>
<th>Sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1C2 15</td>
<td>Chinese – 11</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>Arabic – 1</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Bengali – 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Japanese – 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Telugu – 1</td>
<td></td>
</tr>
<tr>
<td>C1W2 16</td>
<td>Chinese – 10</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>Farsi – 2</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Spanish – 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cebuano – 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Korean – 1</td>
<td></td>
</tr>
</tbody>
</table>

Table 8.2 The percentage of each training session completed by the number of participants in each training group.
8.2.2 Stimuli

The training and testing stimuli and procedures were the same as those used in Experiment Three, with the following exceptions.

8.2.2.1 Training Programs

The reaction time test was presented on DMDX software. As discussed previously, the word-spotting training is also run on DMDX software. In order to avoid a confounding variable of experience gained in using the software, a DMDX component was incorporated into the training in connected speech processes.

The instructions and recordings for the exercises in the connected speech processes training were programmed to run on the DMDX software. This enabled participants in the connected speech processes training group to gain the same types of experiences necessary to complete the word-spotting task and the reaction time post-tests, i.e. sitting at the computer to complete the training, wearing headphones, running DMDX, reading instructions on the screen, hearing a recording, making a response by typing on the keyboard and pressing the space bar to move to the next trial.

The informational content of the training was again presented by the author, but with participants now subsequently turning to their computers to hear example audio files and complete training exercises.

8.2.2.2 Pre- and Post-tests

The pre- and post-test dictations and daily mini-dictations comprised the same sentences as Experiment Three (see Appendices J and K). However, the sentences were randomised prior to commencement of Experiment Four to avoid potential design flaws associated with presenting them in the order in which they were created.
8.2.2.3 Reaction Time Post-test

Upon completion of training session 1, participants were taught 20 low frequency English words which an intermediate student of English as a second language could not reasonably be expected to know. The following reaction time post-test consisted of sentences containing the target words. These sentences did not carry contextual cues to the target word. Ten other sentences that did not contain a target word were included in the reaction time post-test.

The words chosen were mono- or disyllabic. Many were chosen from the Merriam-Webster Word of the Day website, http://www.merriam-webster.com/word-of-the-day/. A full list of the words chosen appears in Table 8.3.

<table>
<thead>
<tr>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>macabre</td>
<td>wisp</td>
</tr>
<tr>
<td>hospice</td>
<td>salvage</td>
</tr>
<tr>
<td>aghast</td>
<td>sachet</td>
</tr>
<tr>
<td>espouse</td>
<td>pageant</td>
</tr>
<tr>
<td>placid</td>
<td>shove</td>
</tr>
<tr>
<td>feisty</td>
<td>turf</td>
</tr>
<tr>
<td>pithy</td>
<td>oblique</td>
</tr>
<tr>
<td>usurp</td>
<td>lanky</td>
</tr>
<tr>
<td>bemused</td>
<td>languid</td>
</tr>
<tr>
<td>wean</td>
<td>stifle</td>
</tr>
</tbody>
</table>

*Table 8.3 Words used in the reaction time post-test at the end of session 1. The day refers to the day taught. All words were presented in the post-test on day 3.*
**Vocabulary Teaching**

The vocabulary teaching sessions were conducted as regular English language classes and taught by the author over two days. Ten words were taught on each day. The post-test was held on a third, following day.

The meanings of the words were taught in addition to the word forms in order to motivate the students to attend the classes, and to permit exposure to the words in a more naturalistic context than by drilling forms.

The lesson materials used can be found in Appendix N. On the first day of vocabulary teaching, the participants were given a handout containing the new words. The handout included a definition stated in simple terms and examples of the words used in sentences. Some definitions were based on those found on the websites of the Cambridge Learner’s Dictionary (http://dictionary.cambridge.org/dictionary/learner-english/) and Cambridge Advanced Learner’s Dictionary (http://dictionary.cambridge.org/dictionary/british/). Information on parts of speech was initially given verbally, but upon participant feedback, was included in subsequent handouts.

Only five words were taught at one time. Emphasis was placed on the Australian English pronunciation of the words, which were repeated often. The vocabulary presentation was followed by a cloze test (in which participants were given sentences missing one newly taught word) to help reinforce the words. Participants completed their cloze tests individually and then checked their answers in groups. Following the teaching of the ten words presented each day, the participants were placed into groups of four and given vocabulary learning activities. Meaningful, group based vocabulary learning activities were used in preference to simply exposing the participants to the words by playing each repeatedly. The activities were run as
games and the winners were given stickers as prizes to motivate involvement. Some of these activities were found in the instructor course book for the Certificate IV in TESOL qualification, published by Totally Trained (2004).

The activities were:

- **Group Story** – groups of participants were randomly given words they must use in a coherent story. The participants worked together to compose the stories, which were then read to the class.

- **Slam** – the new words were presented on cards, which were scattered on the desks in front of each group of participants. Hints relating to each word were given by the author and the participants attempted to be the first in their group to slam their hand down on that word. When the word had been chosen the whole class repeated it in unison. The winner in each group was the participant with the most cards at the end of the game.

- **Word Game** – participants took turns to lift a word card off the top of a deck and explain it to others. If they didn’t know it or the others didn’t guess it, it went to bottom of pile. Both the participant who explained the word and the participant who correctly guessed the answer received a point. The game continued until all the words were used.

- **Synonym or Antonym** - Each group was asked to think of either a synonym or antonym for each word. The team which finished first with correct answers gained a prize. The winning answers were read to the class with an emphasis on frequent repetition of the target word.

- **Speed Game** – participants were put into two teams. One member from each team sat on a chair with their back to the whiteboard and their team mates stood in line facing them. The author wrote one of the newly taught words on
the board and the first two team members in each line tried to explain the meaning of the word without using it. The team that first guessed the word correctly earned their team a point and the game roles were rotated within teams.

**Post-test**

On the third day following the two days of vocabulary teaching, the participants sat at individual laptops and received a reaction time post-test. The participants heard thirty sentences, twenty of which contained target words. Participants were asked to press the right shift key on the laptop when they heard any of the target vocabulary items presented in the sentences. The reaction times of their responses were measured.

**8.2.3 Equipment**

The equipment used in this experiment was the same as that used in Experiment Three.

**8.2.4 Procedure**

The training was conducted for nine days over two weeks. On the tenth day of each session, a dictation post-test was held. The vocabulary was taught over the first two days of the following week, and the reaction time post-test was given on the third day. A one week break was given between training sessions. The schedule can be seen in Table 8.4.
Table 8.4 Training schedule for Experiment Four

<table>
<thead>
<tr>
<th>Week 1</th>
<th>Dictation Pre-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group WiC2</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Group C1W2</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Weeks 2 - 3</strong></td>
<td>Word-Spotting</td>
</tr>
<tr>
<td><strong>Week 3</strong></td>
<td>Dictation Post-Test 1</td>
</tr>
<tr>
<td><strong>Week 4</strong></td>
<td>Vocabulary Teaching and Reaction Time Post-test</td>
</tr>
<tr>
<td><strong>Week 5</strong></td>
<td>No Training</td>
</tr>
<tr>
<td><strong>Weeks 6 -7</strong></td>
<td>CSP</td>
</tr>
<tr>
<td><strong>Week 7</strong></td>
<td>Dictation Post-Test 2</td>
</tr>
</tbody>
</table>

8.3 Results

Unless otherwise specified, the results were analysed using a repeated measures analysis of variance (ANOVA) with training group as the between-groups factor, and an alpha level of .05. Assumptions of normality and homogeneity were satisfactory unless stated otherwise.

Only significant statistics will be reported in this chapter. For full details of the statistical data, see Appendix O.

8.3.1 Dictations

8.3.1.1 Pre-test

Data screening on the partial dictation component of the pre-test revealed that the assumption of normality for the WiC2 group was not met, thus a Mann-Whitney nonparametric test was conducted on the results instead of an independent t test. There was no statistically significant difference between the groups' performance on the set of partial dictations. Assumptions of normality were met for the whole dictation data. An independent t test found there was also no significant difference
between the means for the two training groups. As Table 8.5 shows, the mean scores were quite similar between training groups and across partial and whole components. The high standard deviations show there was wide range of scores in each group.

<table>
<thead>
<tr>
<th>Training Group</th>
<th>Partial</th>
<th>Whole</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1C2</td>
<td>58.80 (13.34)</td>
<td>58.33 (12.08)</td>
</tr>
<tr>
<td>C1W2</td>
<td>60.19 (18.16)</td>
<td>60.94 (13.80)</td>
</tr>
</tbody>
</table>

*Table 8.5 Mean pre-test dictation scores of groups W1C2 and C1W2 as a function of test component (with standard deviations in parentheses). Note N = 31*

*Missegmentations*

Data screening of the missegmentation scores revealed that one participant, LS11, with a z score of 3.36 in the partial dictations, was an outlier, who made substantially more missegmentations than other participants. This participant was removed from further analyses of the missegmentations in the partial dictations and subsequently, normality assumptions were met. The two training groups now contained equal numbers, with 15 participants in each group.

Independent t tests showed there was no significant difference between the two groups on the number of missegmentations in each section of the pre-test. Table 8.6 shows the means for each group.
Pre-test Component

<table>
<thead>
<tr>
<th>Training Group</th>
<th>Partial</th>
<th>Whole</th>
</tr>
</thead>
<tbody>
<tr>
<td>W1C2</td>
<td>6.53 (3.02)</td>
<td>7.13 (3.68)</td>
</tr>
<tr>
<td>C1W2</td>
<td>6.07 (3.33)</td>
<td>7.75 (3.96)</td>
</tr>
</tbody>
</table>

Table 8.6 Mean pre-test missegmentation scores of groups W1C2 and C1W2 as a function of test component (with standard deviations in parentheses).

Correlations

With the one outlier returned to the data, Pearson product-moment correlations were performed between the dictation and missegmentation scores. Again, the two scores were not independent. The aim of the analysis was to establish how strong an impact missegmentation may have on dictation scores.

A strong, negative relationship between partial dictation scores and partial dictation missegmentation scores was significant, $r(31) = -0.59$, $p < .001$, as was the strong negative correlation between whole dictation scores and whole dictation missegmentation scores, $r(31) = -0.61$, $p < .001$. These results clearly show the higher a participant’s dictation score, the lower their missegmentation score.

8.3.1.2 Pre-test – Post-test 1

An analysis was conducted on the scores of those participants who had attended 78% or more of the training: 12 participants in Group W1C2 (10 female, 2 male) and 15 participants in Group C1W2 (11 female, 4 male). The first languages of W1C2 were Chinese (9), Japanese (1), Arabic (1) and Telugu (1). In C1W2 the first languages were Chinese (9), Farsi (2), Spanish (2), Cebuano (1) and Korean (1).
Consistency between the tests was good. The average number of words in the pre-test partial dictations was 6.8, while the average in the post-test 1 partial dictations was 6.4. The average number of words in the pre-test whole dictations was 11 and the average number of words in the post-test 1 whole dictations was 12.

The reported analyses are divided into partial and whole dictation results. The main effect of ‘test number’ in the partial dictations was significant, $F(1, 26) = 99.10, p = <0.001$. Partial dictation scores in post-test 1 were significantly higher than in the pre-test (pre-test: $M = 59.54, SD = 3.26$; post-test 1: $M = 73.99, SD = 2.77$). However, the main effect of ‘training group’ and interaction of ‘training group’ by ‘test number’ were not significant. The analysis of the whole dictation scores showed no significant main effects or interactions. The mean percentage of correct scores is displayed in Figure 8.1. Note that error bars are not included in figures for this experiment as the number of participants in each group, in each session, were quite different.

![Figure 8.1](image_url)
**Missegmentations**

The mean number of missegmentations made by both groups decreased considerably from the pre-test to post-test 1 and for the partial dictations the main effect of test number was significant, $F(1, 26) = 26.12, p < .001$. The main effect of ‘training group’ and the ‘training group’ by ‘test number’ interaction were not significant. The analysis was repeated for the number of missegmentations made in the whole dictations in the pre-test and post-test 1. Here there were no significant effects or interactions. Figure 8.2 shows the number of missegmentations made by each group across the pre-test and post-test 1 in both the partial and whole dictation sets.

![Figure 8.2 Mean number of missegmentations made by each training group over the pre-test and post-test 1, as a function of dictation test type. N=28](image)

**Correlations**

Pearson product-moment correlations were performed between the post-test 1 dictation and missegmentation scores. As in the pre-test analysis, a strong, negative relationship between partial dictation scores and partial dictation missegmentation scores was significant, $r(27) = -.57, p < .001$. A negative correlation between whole dictation scores and whole dictation missegmentation scores almost reached
significance despite the low number of participants and items, $r(27) = -0.38$, $p = .051$.

8.3.1.3 Pre-test – Post-test 2

Structural consistency between the pre-test and post-test 2 was good. Six participants in the C1W2 Group, but only one participant in the W1C2 Group, completed more than 78% of the training and post-test 2. For this reason, the analyses were conducted on the scores of participants who completed 67% or more of the training and post-test 2, giving a total of seven participants in the C1W2 Group and three participants in the W1C2 Group. The language backgrounds of C1W2 were Chinese (5), Farsi (1) and Spanish (1), and in W1C2 they were Chinese (2) and Japanese (1).

In the partial dictations, there was a significant main effect of ‘test number’, $F(1, 9) = 6.37, p = .04$, with the performance of both training groups improving. The mean score of the groups combined was 57% in the pre-test ($SD = 6.30$) and 63% in post-test 2 ($SD = 5.06$). However, there was no significant main effect of ‘training group’ or ‘training group by test number’ interaction. An analysis of the whole dictations combined shows no significant effects or interactions. The mean dictation scores for each test type for each group (composed of the participants who completed 67% or more of the training) are displayed in Figure 8.3.
CH. 8: Experiment Four

Figure 8.3 Mean pre-test and post-test 2 scores gained by the two training groups, as a function of component. In the whole dictations, only the score for Group C1W2, which received word-spotting training, improved. N=10

**Missegmentations**

The mean number of missegmentations decreased numerically from the pre-test to post-test 2 in both the partial and whole dictation sets, as can be seen in Figure 8.4. However, there were no significant effects or interactions for the partial or whole dictation components.

Figure 8.4. Mean number of missegmentations made in each training group in the pre-test and post-test 2, as a function of dictation test type. N=10
Correlations

Pearson product-moment correlations were performed on the post-test 2 dictation and missegmentation data. The negative relationship between the both sets of dictation scores and corresponding missegmentation scores was not significant.

8.3.2 Mini-Dictations

8.3.2.1 Partial Dictations

Initial screening of the partial dictation component of the daily mini-dictation data showed one outlier in Group C1W2. Removing this outlier resulted in 8 participants in Group C1W2 and 4 participants in Group W1C2 who had completed all nine training sessions. Analysing the results of the first six training sessions only (with the one outlier removed) resulted in a larger and more evenly distributed number of participants in each group: 12 participants in Group C1W2 and 10 in Group W1C2.

The main effect of ‘test number’ was significant, $F(5, 21) = 20.51$, $p < .001$. Pairwise comparisons can be found in Appendix O. There was no significant main effect of ‘training group’ or interaction of ‘training group’ by ‘test number’.

As Figure 8.5 shows, the mean partial dictation scores of the two training groups rose and fell over the course of the first six mini-dictation tests. The scores of the two training groups move in the same direction across each test. Of note is the increase in the mean score of 12% in test 4 for those participants who received word-spotting training, placing the group above those who received training in connected speech processes.
Figure 8.5 Mean partial dictation scores for each training group across the first six mini-dictations in session one. The scores of each training group follow the same pattern, suggesting that test difficulty played a strong role. N=22

Missegmentations

Figure 8.6 shows the mean number of missegmentations made by the two training groups in the partial dictations of the first six mini-dictation tests. The results show there was no consistent improvement or worsening of scores for either group.

Figure 8.6 The mean number of missegmentations made in each training group across the partial dictations of the first six mini-dictations in session one. The lower the score, the better the performance. The results do not reflect a clear pattern across groups or advantage for one group. N=22
The main effect of 'test number' was significant, $F(5, 22) = 2.93, p = .02$. Pairwise comparisons showed the only significant difference was between Test 1 and Test 3, $p = .01$, with Test 1 having fewer missegmentations ($M = .93$) than Test 3 ($M = 2.25$). The results were due to a higher number of missegmentations across participants in both training groups and were not merely an artifact of one participant's results. There were no other significant effects or interactions.

**Correlations**

Pearson's product-moment correlations ($r$) were performed between the dictation scores and missegmentation scores for each mini-dictation test. An analysis of the partial dictation component, showed strong significant negative correlations between the mean percentage of correct words in each dictation and the number of missegmentations made for all but two tests. Test 3 was close to significance at the 0.05 level, $p = .066$.

<table>
<thead>
<tr>
<th>Test</th>
<th>Pearson Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>-.322</td>
</tr>
<tr>
<td>Test 2</td>
<td>-.590 **</td>
</tr>
<tr>
<td>Test 3</td>
<td>-.389</td>
</tr>
<tr>
<td>Test 4</td>
<td>-.838 **</td>
</tr>
<tr>
<td>Test 5</td>
<td>-.499 *</td>
</tr>
<tr>
<td>Test 6</td>
<td>-.860 **</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

*Table 8.7* Correlations between the percentage of correct words and number of words missegmented in the partial dictation component of the daily mini-dictation tests. $N=22$

### 8.3.2.2 Whole Dictations

Data screening showed there were no outliers in the whole-dictation data. This allowed all 23 participants (13 in Group C1W2 and 10 in Group W1C2) who completed
the first six training sessions to be included in the analysis. An analysis was conducted on the mean results for each group across the set of whole-dictations for those six tests. Figure 8.7 shows the mean scores for each test across each training group.

![Figure 8.7 Mean whole dictation scores for each training group across the first six mini-dictations of session one. Consistency in direction of change between the groups suggests test difficulty was again a major determinant of result. N=23](image)

As Figure 8.7 shows, the mean scores of the two training groups over the first six mini-dictation tests were very close. The main effect of ‘test number’ was significant, \( F(5, 22) = 20.26, p < 0.001 \). The main effect of ‘training group’ and ‘test number’ by ‘training group’ interaction were not significant. The significant pairwise comparisons can be found in Appendix O.

**Missegmentations**

The mean number of missegmentations made by both training groups in the whole dictation component of the daily mini-dictations was also analysed. Figure 8.8 displays the number of missegmentations made by each training group in the whole dictation component of each mini-dictation. With the exception of an increase for...
C1W2 (which received training in connected speech processes in this training session) in Test 2, both groups improved in the number of missegmentations made until the fifth and sixth test, in which substantially more missegmentations were made.

Mauchly’s test of sphericity indicated that the sphericity assumption was not met for ‘test number’. The main effect of ‘test number’ was significant with a Huynh-Feldt adjustment to the degrees of freedom, $F(4.06, 85.33) = 12.59, p < .001$. The main effect of ‘training group’ was not significant.

In this analysis, the ‘training group’ by ‘test number’ interaction was also significant, $F(4.06, 85.33) = 2.60, p = .04$. This interaction was analysed by means of simple effect analyses, which were performed separately for the two training groups. Using the Huynh-Feldt correction, the simple effect for Group W1C2 was significant, $F(4.32, 38.91) = 6.60, p = < .001$. The simple effect for Group C1W2 was also significant, $F(3.65, 43.84) = 7.96, p = < .001$.

Each simple effect analysis was followed by simple contrasts comparing the second to sixth tests to the first test. The simple comparisons showed that Group W1C2, which
received word-spotting training, made significantly more missegmentations in Test 6 than in Test 1, $F(1, 9) = 6.98, p = .03$. This was led by one participant who made 12 missegmentations in Test 6, but only one in Test 1. Group C1W2, which received training in connected speech processes in session 1, made significantly more missegmentations in Test 5, $F(1, 12) = 14.90, p = .002$, and Test 6, $F(1, 12) = 6.94, p = .02$, than they had in Test 1. However, they also made significantly less missegmentations in Test 4 than they had in Test 1, $F(1, 12) = 6.31, p = .03$, (Test 1 $M = 1.69$, Test 4 $M = 0.85$).

**Correlations**

Pearson’s product-moment correlation analyses on the whole dictation component of the tests produced only one significant negative correlation between the dictation scores and missegmentation scores, Test 1, $p = .001$. Analyses of all other tests showed negative correlations that were not significant. Repeating the analyses for each training group separately produced the same pattern of results.

### 8.3.3 Word-Spotting Scores

#### 8.3.3.1 Mean Percentage Correct

The two groups were initially combined and an analysis performed on the daily results achieved in the word-spotting training task. The first five days of training were analysed, giving a total of 21 participants (8 in Group W1C2, which received word-spotting training in session 1, and 13 in Group C1W2, which received word-spotting training after an initial program of connected speech processes training). In the combined groups analysis, the main effect of ‘test number’ was significant, $F(4, 20) = 9.22, p = <.001$. Pairwise comparisons show a significant difference between Test 4 and all other tests, with the mean score of Test 4 being higher than that of each other group. This can be explained by the training presenting repeated stimuli,
recorded by different speakers, from day 4 on. The scores achieved on Day 5 should be comparable for the same reason, but they are considerably lower. Mean scores across the six testing times are shown in Figure 8.9.

\[ \text{Figure 8.9 Mean scores achieved in the word-spotting task. The peak in scores on day 4 may be due to the repetition of stimuli, recorded by different speakers. N=21} \]

The combined word-spotting scores do not improve with repeated exposure. However, a comparison of the two training groups reveals a significant difference in their performance. The mean effect of ‘test number’ was significant, \( F(4, 20) = 8.59, p < .001 \), as was the interaction between ‘test number’ and ‘training group’, \( F(4, 20) = 2.71, p = .036 \). The mean effect of ‘training group’ was not significant.

The ‘test number’ by ‘training group’ interaction was analysed by means of simple effect analyses performed separately for the two training groups using separate error terms. Each simple effect analysis was followed by simple contrasts comparing each level of test number to the pretest. The simple effect for Group W1C2 was significant, \( F(1, 7) = 4.44, p = .007 \), with the simple comparisons showing that the group’s mean score on Day 4 was significantly higher than it was on Day 1, \( F(1, 7) = 7.81, p = .027 \).
The simple effect for Group C1W2 was also significant, $F(1, 12) = 8.19, p < 0.001$, with the simple comparisons showing a significantly lower mean score on Day 3 than had been gained on Day 1, $F(1, 12) = 6.86, p = .022$. Group C1W2 had particular difficulty with the third day of training. Here the decrease in mean scores reflects a drop in scores of 11 of the 12 participants in the C1W2 group from Day 2 to Day 3.

Figure 8.10 shows that Group W1C2 did improve over the first four training days, while the mean scores of Group C1W2 reduced over the first three training days before rising on the fourth. The scores of both training groups fell on the fifth day. It is likely the rise in scores on day 4 can be explained by the repetition of stimuli, recorded by different speakers, from day 4.

8.3.3.2 Mean Percentage Segmented Correctly

The previous analyses on the word-spotting data of Experiment Three and Experiment Four were conducted on the percentage of correct responses. This may be problematic as word selection can be influenced by simple errors, such as

![Bar chart showing mean percentage correct by group and day](image-url)
misperceptions of consonants (especially on the voicing feature) and vowels, which, if occurring word-externally, are irrelevant to speech segmentation. In acknowledgment of this fact, in this experiment the data of each training group was scored again, this time for correct segmentation of the string.

An analysis was conducted on the segmentation data from the first five days of the training for the same participants as in the previous analyses. An analysis of the training groups combined resulted in a significant main effect of ‘test number’, $F(4, 20) = 19.37, p < .001$. Figure 8.11 shows these combined results.

![Figure 8.11 The mean percentage of correctly segmented word-spotting items by all participants across five testing times. N=21](image)

**Correlations**

Pearson product-moment correlations were performed between the percentages of items segmented correctly in the word-spotting task and the daily missegmentation scores for each participant. Again, there was a strong significant correlation between the word-spotting and missegmentation scores on Day 1, $r(10) = -.77, p = .010$, which showed that the better the participants segmented the word-spotting material, the lower the number of missegmentations made in the first mini-dictation. There were
negative correlations between the data from all further days, but none reached significance.

8.3.4 Reaction Time Post-test

An independent $t$ test was conducted on the mean scores for each participant in the reaction time post-test of session 1. Only the scores of participants who had received 78% or more of the training (no more than two missed training days) and had attended both vocabulary teaching sessions were included - 15 participants in Group C1W2 and 9 participants in Group W1C2.

The mean reaction time was numerically lower in Group W1C2, which had received word-spotting training before the vocabulary teaching and RT test ($M = 1.78 \text{ s}, SD = 0.59 \text{ s}$), than Group C1W2, which had received training in connected speech processes ($M = 2.11 \text{ s}, SD = 1.02 \text{ s}$). However, the results did not reach significance.

It should be noted here that during the vocabulary training, the Spanish speaking participants remarked upon the similarity between some of the English words taught and words in their L1. This did not appear to afford them an advantage, in that the C1W2 group, which contained both Spanish speakers, obtained slower reaction times than the W1C2 group.

8.3.5 Questionnaire

The questionnaire given to participants in both Experiments Three and Four can be found in Appendix M. 12 participants completed the questionnaire at the end of session 2. The number of participants who rated each training group as helpful or not helpful is shown in Table 8.8.
Table 8.8 Questionnaire results. The number of questionnaire respondents who rated each training type as most and least helpful. N=12

<table>
<thead>
<tr>
<th></th>
<th>Most Helpful</th>
<th>Least Helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>CSP</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>No response</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

Again, the training in connected speech processes was seen as being more helpful than the word-spotting training. As discussed previously, this may be due, at least in part, to the participants’ greater respect for and belief in teacher-led, explicit lessons.

The participants’ responses to the training were very similar to those of participants in Experiment Three. Many of the participants who rated the training in connected speech processes as the most helpful commented favourably on being explicitly taught the pronunciation of Australian English, particularly word stress. Most comments related to enhanced listening skills, for instance, the recognition of speech that is hard to distinguish, but some also mentioned the value of being able to practice pronunciation.

However, although the program taught receptive, rather than productive skills, one participant criticised the training in connected speech processes as they believed it was not genuine listening training because it focused on pronunciation. The other two participants who rated the training as least helpful stated that they felt the program was too short.

Of those who rated the word-spotting training as the most helpful, two stated it was because it provided training in vowel recognition. The third wrote that it was because it was more interesting, so they were motivated to pay attention and learn.
Three participants who had rated training in connected speech processes as being most helpful stressed that the word-spotting training was also beneficial. All commented that the training was helpful in recognising the correct pronunciation of words.

It is clear that, despite being given the rationale for the training in the information session, the participants felt the value of the word-spotting training lay in its provision of practice in the recognition of words, specifically through the improvement of one’s knowledge and perception of Australian English vowels. The concept of speech segmentation seems to be a difficult one to grasp. Nonetheless, the participants are correct that the training does have potential for use as a method of training vowel recognition. Future analyses of the current data may show a training benefit in this regard.

The word-spotting training was criticised by one participant as not improving listening skills, which they had defined as “understanding a sentence or a paragraph”. The repetition involved in the training was also criticised. Another complaint was that the nonsense string often sounded like a real word. It is likely that the use of reduced vowels in the sets targeting perception of stressed onsets posed a problem for listeners. This issue could be expected to be resolved with on-going practice and strengthens the value of the training as a future method of training vowel recognition.

The descriptive adjectives chosen by the participants to describe the two training methods are shown in Figure 8.13. Of particular note is the perception that the word spotting training was long and difficult and the training in connected speech processes was short and easy.
8.4 Discussion

Once again word-spotting training was shown to be at least as effective in improving second language listening as training in connected speech processes, an approach which has previously been found to be effective (Brown & Hilferty, 1986; Kennedy & Blanchet, 2013). There was no significant difference between the training groups in the post-tests. Both groups performed very similarly, significantly improving in most partial dictation measures, with no significant results on whole dictation measures.

It is possible that both training methods were effective in reducing the occurrence of missegmentations in the post-tests, with the number of missegmentations made in
the partial and whole components of the two post-tests significantly decreasing. In post-test 1, the mean number of missegmentations made by both groups decreased quite considerably and there was a strong significant effect of test number for the partial dictation components.

In the mini-dictations, with the exception of a drop from Day 1 to Day 2, there was an increase in the mean percentage correct for both groups to a peak on Day 4. Both groups also improved in the number of missegmentations made in the mini-dictations until the fifth and sixth test, except for an increase for C1W2 (which had received training in connected speech processes) in Test 2. Group C1W2, which had received training in connected speech processes, made significantly less missegmentations in Test 4 than they had in Test 1. A correlation analysis between the word-spotting and daily missegmentation scores from the mini-dictation data again showed a strong negative relationship on Day 1.

The results of this experiment also underline the importance of segmentation skills. It was again shown that there is generally a very strong negative correlation between dictation scores and the number of missegmentations made. In other words, it can be suggested that the higher the number of missegmentations made, the lower the dictation score achieved.

Most importantly, the word-spotting task itself was shown to result in statistically significant improvements to speech segmentation performance when analysed as the percentage of items correctly segmented per set. Therefore, it seems that the word-spotting training is doing what it was designed to do - improve segmentation skills.
9.1 Aim, Design and Hypotheses

The aim of Experiment Five was to gain data that was, in certain respects, more reliable than previous experiments had allowed. This was achieved through conducting the training over a shorter period of time in a laboratory setting with a pre-determined minimum number of participants and the use of an established psycholinguistic tool as the post-test.

The study followed a treatment – post-test design. In this experiment, participants received only one type of training, either training in the word-spotting method or in connected speech processes. Each training method was delivered at individual computers over three days.

The post-test used was an Artificial Language Learning (ALL) task. This task allows considerable control over the testing environment. In this respect it contrasts with the dictation post-tests used in the preceding experiments, despite the widespread past use and face validity of dictation as a testing tool. There are many variables involved in both the design of dictations (e.g., in test complexity across tests) and also
the performance of the tests (e.g., mistakes are made for reasons other than missegmentation).

Artificial languages composed of phonemes found in natural languages have been used in numerous studies of adult segmentation ability (for instance, Saffran et al., 1996b; Saffran, Newport, Aslin, Tunick & Barrueco, 1997; Fernandes, Kolinsky & Ventura, 2009; Tyler & Cutler, 2009; Hoch, Tyler & Tillman, 2013). In the task participants are exposed to an artificial, synthesised language which concatenates meaningless sequences and is presented as a continuous sequence of speech. Participants are then tested on their ability to recognise the component ‘words’ of the language.

ALL studies are similar in the procedures used. The seminal study was conducted by Saffran et al., (1996b). Adult participants were exposed to a continuous stream of artificial trisyllabic CV words to study whether transitional probabilities were used as cues to word boundaries by adult listeners. In a second experiment a vowel lengthening cue was included. Participants were given one of two versions of forced choice tests at the end of the 21 minute exposure period. One alternative set words from the artificial language against words composed of novel sequences of CV pairs found in the language. The other presented artificial words alongside part-words consisting of two syllables that occurred in sequence within a word joined to another syllable.

More recently, Hoch et al. (2013) compared statistical learning of verbal versus nonverbal materials, and of words of regular versus irregular lengths. Native English speaking adults were exposed to one of four conditions: 1) verbal languages with units of regular length, 2) verbal languages with units of irregular length, 3) nonverbal languages with units of regular length, and 4) nonverbal languages with units of
irregular length. The participants displayed better performance in learning verbal than non-verbal materials, and in learning words with regular lengths than words with irregular lengths.

The materials used in the present study were based on those used in the second condition of the Hoch, Tyler and Tillman (2013) study. In that condition, the familiarity sequence consisted of a continuous stream of artificial two and three syllable words composed of CV strings. The exposure period was just over 12 minutes long, separated into five equal segments. It was followed by a two-alternative forced choice test of 24 word pairs, composed of a word from the artificial language and a part-word, which consisted of one or two syllables from one word and one syllable from another. In this way, each part-word contained a word-internal word boundary.

In the present experiment, artificial words of two, three and four syllables were presented as continuous streams of synthesised speech. Left-edge pitch movement was used as a word boundary cue to mirror the predominance of initial stress in English words.

Training was given to students (undergraduate and postgraduate) at the University of Western Sydney. There were two different participant groups in this study:

1) monolingual English speaking students (L1)

2) students who spoke English as an additional language and who had been in residence in Australia for less than two years (L2). The groups will be referred to henceforth as L1 and L2, respectively. The L1 group was included as a control comparison for the L2 group.
The hypotheses were:

1) Participants who had received word-spotting training would perform at least as well at the ALL task as those who had received training in connected speech processes, which has been shown in Kennedy and Blanchet (2013) to improve perception reliant on speech segmentation;

2) The word-spotting training, which indirectly provides practice in segmenting strings before initial stressed syllables, might lead to higher ALL task scores than the training in connected speech processes, which, while also teaching word stress, places greater focus on word-internal processes.

3) The L1 group would receive higher ALL task scores than the L2 group as left-edge prominence is a native English speech segmentation cue.

4) The L1 group would perform significantly better than the L2 group at the word-spotting task, but the L2 group would show greater improvement over the three training sets than the L1 group.

5) Following on from Experiment Four results, the word-spotting responses would show greater improvement across sets when analysed as the percentage of items segmented correctly than they would when analysed as the percentage of correct responses given.

9.2 Method

9.2.1 Participants

Participants were 48 students of the University of Western Sydney. 24 of these were born in Australia, spoke only English at home and did not have significant exposure to a second language (had not received instruction in a foreign language beyond high
school and had never lived in a country where a foreign language was spoken). The other 24 participants spoke English as an additional language and had lived in Australia for less than two years.

Six participants from each of the L1 and L2 groups were randomly assigned to each of four groups. The ALL task results were counter-balanced across the dimensions of language, training received and ALL exposure set. The details of each group are summarised in Table 9.1.

<table>
<thead>
<tr>
<th>Number of participants</th>
<th>Language Background</th>
<th>Training Received</th>
<th>ALL Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>L1</td>
<td>WS</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>L2</td>
<td>WS</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>L1</td>
<td>CSP</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>L2</td>
<td>CSP</td>
<td>A</td>
</tr>
<tr>
<td>6</td>
<td>L1</td>
<td>WS</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>L2</td>
<td>WS</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>L1</td>
<td>CSP</td>
<td>B</td>
</tr>
<tr>
<td>6</td>
<td>L2</td>
<td>CSP</td>
<td>B</td>
</tr>
</tbody>
</table>

*Table 9.1 The composition of the four Training x Exposure groups.*

As can be seen in Table 9.2, although the assignment of L2 participants to groups was randomised, the balance of first languages between the groups was very good.
Table 9.2 The language backgrounds of the participants in each Training x Exposure group. N = 48

<table>
<thead>
<tr>
<th>Training X Exposure</th>
<th>Number of participants</th>
<th>Language Background</th>
</tr>
</thead>
<tbody>
<tr>
<td>WS-A</td>
<td>4</td>
<td>Chinese</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hindi</td>
</tr>
<tr>
<td>CSP-A</td>
<td>3</td>
<td>Chinese</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bengali</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urdu</td>
</tr>
<tr>
<td>WS-B</td>
<td>3</td>
<td>Chinese</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pashto</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Russian</td>
</tr>
<tr>
<td>CSP-B</td>
<td>4</td>
<td>Chinese</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dutch</td>
</tr>
</tbody>
</table>

Table 9.3 presents the mean ages of all participants in each independent variable. The balance here was also very good.

<table>
<thead>
<tr>
<th>Language group</th>
<th>Training received</th>
<th>ALL exposure</th>
<th>Mean Age</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L1</td>
<td>L2</td>
<td>WS</td>
</tr>
<tr>
<td>Mean Age</td>
<td>23 (3.80)</td>
<td>25 (3.99)</td>
<td>24 (4.55)</td>
</tr>
</tbody>
</table>

Table 9.3 Mean ages of participants (with standard deviations in parentheses). N=48

Table 9.4 displays the mean length of residence (LOR) in Australia for L2 listeners within each of the Training X Exposure groups. The average length of residency was lower for participants receiving word-spotting training than that of participants receiving training in connected speech processes.
The study was advertised to first year psychology students on the SONA system, by group email to all students and an additional email to international students only. Flyers were placed around the university and information was given in lectures run by the Department of Linguistics.

First year psychology students received course credit for participating. All other students were paid $60 if they completed all three days of training and the following ALL post-test.

9.2.2 Stimuli

9.2.2.1 Training Programs

The training programs run in Experiment Five were again the word-spotting and connected speech processes training. This time each training program was offered over only three days.

Only the first three sets of the word-spotting training were offered: sets A01, A02 and A03. As a result, there was no repetition of word-spotting items during this experiment.

The shortened training in connected speech processes contained lessons on syllable stress, reduced vowels in function words, flaps, -nt reduction and glottal stops. As
training was conducted by each participant individually in a self-guided format, the
teaching component of the training in connected speech processes could not be given
by the author. Lesson sheets were given to each participant taking part in the
training program on each day. These can be found in Appendix P (note that the
training in connected speech processes was again delivered under the title Sound
System Awareness). The handouts were based on Hagen (2000). Participants
practiced each language point featured in the training by performing the same DMDX
sets presented in earlier experiments and also completing dictations on the lesson
sheet.

9.2.2.2 Post-test

Two artificial languages were constructed to serve as different exposure sets and post-
test items. The languages will be referred to as A and B, respectively. Each language
contained 6 ‘words’ (listed in Table 9.5); two with two syllables, two with three
syllables and two with four.

The words were composed of CV sequences, formed through a combination of the
consonants /b, d, m, n, f, s/ and the vowels /a, i, u/. The vowels appear in all of the
first languages of the participants. The consonants appear in most of the languages,
with the exception that standard Mandarin (which I use here for simplicity) contains
voiceless, but not voiced stops and the Russian language contains dental versions of
the alveolar consonants, /d/, /n/ and /s/ (Maddieson, 1984). These one-feature
differences were not considered likely to cause differences in task performance.

The syllable sequences of words in one language formed part-words in the other; that is, the first syllable of disyllabic words in Language A form the final syllable of

---

16 I am grateful to Michael Tyler for creating the ALL materials and providing detailed notes on their construction.
corresponding words in Language B and vice versa, the first two syllables of trisyllabic words in Language A form the final two syllables of corresponding words in Language B, and the first three syllables of tetrasyllabic words in Language B form the final three syllables of corresponding words in Language A. Importantly, no syllable of any word of either language was repeated within that language.

<table>
<thead>
<tr>
<th>Language A</th>
<th>Language B</th>
</tr>
</thead>
<tbody>
<tr>
<td>'safa</td>
<td>'fabu</td>
</tr>
<tr>
<td>'busi</td>
<td>'sisa</td>
</tr>
<tr>
<td>'nafiba</td>
<td>'nunafi</td>
</tr>
<tr>
<td>'midanu</td>
<td>'bamida</td>
</tr>
<tr>
<td>'sumadifu</td>
<td>'madifuni</td>
</tr>
<tr>
<td>'nimudubi</td>
<td>'mudubisu</td>
</tr>
</tbody>
</table>

Table 9.5 The ‘words’ of each artificial language.

The words of each language were synthesised with the MBROLA text-to-speech software (Dutoit, Pagel, Pierret, Bataille, & Van der Vrecken, 1996), using the French female (fr2) diphone database. For each language, the words were then concatenated in a pseudorandom order to create roughly 11 minutes of continuous artificial speech, forming separate exposure material for the two languages. To make it easier for participants to sustain attention, the exposure sets were divided into five blocks of two minutes and 12 seconds’ duration. Each word was repeated 32 times per block, making 160 repetitions of each word across each exposure phase.
Every word designed for use in the exposure phase was created with a left-edge pitch movement cue to approximate a stressed syllable. Fundamental frequency (F0) on initial syllables fell from 272Hz to 220 on following syllables. (Words contrasted in the test phase had a constant F0 of 220Hz across all syllables.) Left-edge pitch movement and the transitional probabilities between syllables were the only cues to word boundaries. Consonant and vowel length was set at an equal duration of 116 ms each (as in Tyler & Cutler, 2009). The intensity of vowels was close to 75dB on all syllables.

While participants heard different exposure sets depending on which group they had been allocated to, they all completed the same two-alternative forced choice test phase; the only difference being that the DMDX files were programmed differently to record responses as either correct or incorrect, depending on the assigned language.

The test phase consisted of 24 pairs of words – one from each exposure language - of equal syllable length. Prosody was held constant across all syllables of the test words. To avoid the potential for order effects, each pair was repeated during the test in the reverse order. The author programmed DMDX to present the five blocks of the exposure phase and the following test phase, randomising the presentation of all test pairs.

**Comparison with the Word-Spotting Material**

There were three similarities between the ALL material and target words in the word-spotting material. First, words in the ALL material contained a rise in F0 on the initial syllable, corresponding to the Stress condition in the word-spotting material. Second, each word in the ALL material was vowel-final and consonant-initial, which corresponded to the Difficult Transition condition in the word-spotting material, in that legal English phonotactic sequences needed to be divided in order to accurately
segment the string. Finally, both stimuli sets also contained words of varying lengths.

The ALL material was also different from the word-spotting material in a number of respects. While the stimuli did share the initial F0 rise of the word-spotting Stress condition, none of its ‘syllables’ contained reduced vowels or increased duration. For this reason, the perception of initial stress was cued in the ALL material by a rise in pitch alone. Another difference is that there were no boundaries in the ALL material that corresponded to the Easy Transition condition in the word-spotting material as none of the words ended in consonants. For this reason, word boundaries were not suggested by sequences of consonants. A further difference between the stimuli sets was that the ALL material contains only CV syllable shapes, whereas the word-spotting material contains complex onsets and codas.

A further important difference between the stimuli sets was that in the ALL task there was no superfluous content in the continuous string. Every syllable of the ALL material was part of a word, whereas the word-spotting material contained monosyllabic nonsense contexts. Therefore, in contrast to the word-spotting task, where one syllable of the string could be quickly ‘discarded’ from short term memory, each syllable of the ALL material was important for the duration of the task.

Exposure to the word-spotting task was thought to have benefit for participants in completing the ALL task, as it provided an abundance of practice in locating word boundaries at initial syllables carrying pitch movement. In contrast, the training in connected speech processes was not expected to provide training benefits for two main reasons. It featured only one session of stress-related lessons, and these were on identifying a range of lexical stress patterns. Certainly, the participants would be exposed to a large number of words with initial stress throughout the training (such is
the ubiquity of the stress patterning in English), but would not receive repeated, controlled practice in segmenting continuous speech at stressed syllables. Second, the majority of the training focused on the perception of reduced forms, which, as discussed above, were not a feature of the ALL material.

9.2.3 Equipment

All the training and testing was conducted using DMDX software, version 3.2.0.1 (Forster & Forster, 2003), with participants wearing Koss headphones and seated at Lenovo ThinkVantage T61p laptops.

9.2.4 Procedure

9.2.4.1 Daily Training Schedule

Participants attended the training in individual timeslots. Both the word-spotting and connected speech processes training sessions were conducted over three consecutive days. The author gave the participants verbal information on how to complete the training and ensured understanding of the task before each training session commenced.

The word-spotting training was run in a self-study format in the same way as in the previous two experiments. Participants received immediate feedback on the correct answer for each item after they had submitted each response. However, participants were not given feedback on the total percentage they answered correct each day as had occurred in previous experiments.

Participants undertaking the training in connected speech processes received printed information each day (Appendix P), containing lesson content, instructions for how to operate the DMDX software to complete practice sessions, and sections in which to
write cloze test and dictation responses. Participants received immediate feedback on whether each response they gave in DMDX trials was correct. However, they did not receive feedback on their answers to cloze tests or dictations.

9.2.4.2 ALL post-test

On the fourth consecutive day of the program, participants completed the ALL task training and test phase. The author explained to each participant that they would hear words from an artificial language presented as a continuous stream, with no pauses between them, and that in the following test phase they would be asked to identify individual words from the language. They heard the exposure phase sound files for either Language A or B depending on which group they had randomly been allocated to. Participants were told they could rest for as long as they needed between exposure blocks and instructions were given on how to operate the DMDX program to move between them.

The initial DMDX screen showed the following instructions: "Relax and listen to the language. When you are ready to start, press the spacebar." At the start of each exposure phase, the number of the block was given, for example “(1st of 5)”, then the instructions, “Listen carefully. When the language stops, press the spacebar”. When the participants had pressed the spacebar after the sound file had stopped, the following lines were displayed, “Please take a short break. Press the spacebar when you are ready to continue”.

At the end of the exposure phase, participants completed the forced choice test phase. As discussed previously, participants heard the same randomised test pairs, regardless of the version of the language they were exposed to. The author explained how participants were to select the correct word from the two alternatives using the
keyboard. They answered by pressing the left shift key (labeled as ‘1’) to choose the first word they heard and the right shift key (labeled as ‘2’) to choose the second.

The first screen of the test phase presented the following lines, "Which word is part of the language?" and the following instructions, "Press 1 or 2. Press spacebar to start". Each word’s sound file was presented with a concurrent visual display of the number “1” or “2”, as appropriate. Sound files were played with a 333ms interstimulus interval between them. After each trial, the instructions: "Press 1 or 2. Press spacebar to continue." were displayed.

9.3 Results

Only significant statistics are reported here. Full details of all results can be found in Appendix Q.

9.3.1 ALL post-test

A 2 X 2 X 2 (training received X language background X ALL exposure) between-groups ANOVA was performed on the ALL test phase scores. There were no significant main effects of training, language, or ALL exposure. The ‘training by language’ interaction was also not significant. Figure 9.1 displays the mean percentage of correct responses in the ALL task by each language group as a function of the training received.
The ‘language by ALL exposure’ interaction was significant, $F(1, 47) = 5.27, p = .027$, but not the ‘training by ALL exposure’ interaction. The interaction of ‘training by language by ALL exposure’ was also not significant. The means for each group are given in Table 9.6.

<table>
<thead>
<tr>
<th>Training Received</th>
<th>Language Background</th>
<th>ALL Exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>WS</td>
<td>57.81</td>
<td>15.75</td>
</tr>
<tr>
<td>CSP</td>
<td>58.68</td>
<td>20.54</td>
</tr>
</tbody>
</table>

*Table 9.6 The mean percentage correct in the ALL task for respective training, language and ALL exposure groups. Note: in each group $N = 24$."

As can be seen in Table 9.6, the mean scores for the two training groups were very similar, showing that the training received did not have an impact on participants’ performance of the ALL task. This result is in line with the hypothesis that participants who received word-spotting training would perform at least as well as
those who received training in connected speech processes, which has been empirically shown to improve speech segmentation by Kennedy and Blanchet (2013).

As also hypothesised, while not a significant result, the mean score for native English listeners was higher than that for participants from a different first language background. The L1 group might have benefited from the left-edge pitch movement cue, which is used by native English listeners to segment speech. The ALL materials are not otherwise English-specific. As discussed previously, they contained basic CV strings and the phonetic segments used did not substantially differ from those found in the first languages of L2 group participants.

The mean scores for the two ALL exposure sets differed slightly, while again not a significant result. The difference in performance of participants receiving different exposure sets was unexpected.

Further analyses of the ‘language by ALL exposure’ interaction showed that the simple effect of language background was significant for ALL exposure set A, $F(1, 23) = 7.97, p = .007$, but not for ALL exposure set B, $F(1, 23) = .16, p = .69$. In addition, the simple effect of ALL exposure set was significant for the L1 group, $F(1, 23) = 6.13, p = .017$, but not for the L2 group, $F(1, 23) = .55, p = .46$.

The data was arranged into four groups depending on the participants’ language background and ALL exposure set received: L1A, L1B, L2A and L2B. A univariate analysis of variance was conducted with Games-Howell post-hoc tests performed (as Levene’s Test of Equality of Error Variances was violated).

Multiple comparisons show that native English listeners performed significantly better at the ALL task after receiving exposure to Language A than native listeners who received exposure to Language B, $p = .018$, with a mean difference of 17.38%. In
addition, after exposure to Language A, native listeners achieved significantly higher scores than participants from the L2 group, $p = .019$, the mean difference being 19.79%. No other pairwise comparisons were significant. Table 9.7 shows the mean scores of each 'Language x ALL Exposure' combination.

<table>
<thead>
<tr>
<th>Language X Exposure</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1A</td>
<td>71.18%</td>
<td>12.87</td>
</tr>
<tr>
<td>L1B</td>
<td>53.88%</td>
<td>13.35</td>
</tr>
<tr>
<td>L2A</td>
<td>51.38%</td>
<td>16.70</td>
</tr>
<tr>
<td>L2B</td>
<td>56.58%</td>
<td>23.60</td>
</tr>
</tbody>
</table>

*Table 9.7 The mean percentage score of each Language x ALL Exposure group. The L1A group performed significantly better than the L1B and L2A groups.*

While native listeners who received exposure to Language A did perform very well at the ALL task, there was no obvious difference between the two exposure languages that would have made Language A easier for native English listeners to segment. The differences observed between groups L1A and L1B were more likely due to the poor performance of two native speakers who had been exposed to Language B and received scores of 37.5% and 29.17%, respectively. The participant who received the lower score had remarked that she had found the ALL exposure blocks as relaxing as meditation. Perhaps, as a result, she had not paid as much attention to the task as other participants.

There was no correlation between the length of residence (LOR) of the L2 group participants and their ALL task scores. Indeed, one of the highest ALL task scores, 83%, was achieved by a participant who had been in Australia for only two months, while a relatively low score of 29% was gained by a participant who had lived in the country for one year.
9.3.1.1 Exploring First Language Group Results

The L2 group scores were coded as belonging to one of three groups, depending on the first language of each participant. The groups were Chinese \((N = 14)\), Indo-Aryan \((N = 4)\) and West-Germanic \((N = 5)\). As there was only one speaker of a Slavic language, the score gained by the Russian participant was excluded from the analysis. There was no correlation between the first language of the participants in the L2 group and their ALL test score. Table 9.8 displays the results.

<table>
<thead>
<tr>
<th>First Language Group</th>
<th>Mean</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese</td>
<td>54.46</td>
<td>17.86</td>
</tr>
<tr>
<td>Indo-Aryan</td>
<td>40.63</td>
<td>20.23</td>
</tr>
<tr>
<td>West-Germanic</td>
<td>63.33</td>
<td>26.58</td>
</tr>
<tr>
<td>cf. English L1</td>
<td>62.50</td>
<td>15.21</td>
</tr>
</tbody>
</table>

*Table 9.8 The mean ALL scores of each group of first languages within the L2 Group. Note that the mean score of the West-Germanic group was the highest of the groups and close to the mean score gained by the L1 Group of native English listeners.*

Pairwise comparisons also showed no significant differences between the groups. However, as expected, the West-Germanic group of Dutch and German listeners benefited from the left-edge prominence cue and gained the highest mean score. As discussed in Chapter 2, Dutch and German, like English, have stress-based rhythm with predominantly initial stress. The mean score received by the listeners in the West-Germanic group was very close to that received by the native listeners of English, itself a West-Germanic language.
**Indo-Aryan Listeners**

The poor performance of the Indo-Aryan group in using pitch movement as a cue to word-initial boundaries is also of interest. Modern Indo-Aryan languages feature lexical stress with predictable fixed stress, but the details vary across languages (Munshi, 2006).

The Indo-Aryan group in this experiment consisted of four participants who spoke Pashto, Hindi, Urdu, and Bengali, respectively. In Pashto, stress is phonemic and variable (Elfenbein, 1997). It is phonetically realised as increased fundamental frequency (F0) (which is perceptually realised as increased pitch) and intensity (Tegey & Robson, 1996) and vowel duration (Bečka, 1969). According to Tegey and Robson (1996) stress is located on the heaviest syllable in a word. It is primarily located in word final position if the word ends in a consonant, and in penultimate position in words ending in a vowel (Tegey & Robson, 1996). As Elfenbein (1997) notes, Pashto stress may also fall on the first syllable. However, in the majority of words, its location is word final, and additional morphological rules increase the potential for stress to be realised in the word final position (Bečka, 1969; Revithiadou, 1999). Due to the predominance of word final stress in their L1, monolingual Pashto listeners would be unlikely to associate pitch movement with word initial boundaries.

According to Baart (2014) stress in Hindi and Urdu is placed on the heaviest syllable in a word, or the last of these if there is more than one candidate. If syllable weight is equal throughout the word, the stress is placed on the penultimate syllable. These facts point to right edge favouring stress placement, at least in words of more than two syllables.

It is the phonetic realisation of stress in Hindi and Urdu that provides the most relevant information on how the participants in the current experiment may have
perceived the left-edge pitch movement boundary cue in the ALL materials.

Empirical research has shown that duration and intensity do not play a role in the realisation of lexical stress in Hindi and Urdu (Dyrud cited in Baart, 2014). Prominence in Hindi and Urdu is signalled by pitch movement. The stressed syllable carries low pitch, which “rises through the syllable boundary” and peaks in the following syllable (Baart, 2014, p. 4). As it is the syllable following the stressed syllable that carries the highest pitch, words with initial stress would not carry the peak of pitch movement on that initial syllable. For this reason, in addition to the tendency for right edge stress placement, monolingual Hindi and Urdu listeners would also be unlikely to associate the peak of pitch movement with word initial boundaries.

In contrast to the other Indo-Aryan languages discussed here, all Bengali-native words confirm to an inviolable rule of initial stress (Chakraborty & Goffman, 2011). Only words borrowed from Sanskrit may violate this rule (Chatterji, 1921). In light of this fact, Bengali listeners might have been expected to benefit from the word initial pitch movement cue presented in the ALL materials in this experiment.

The results of the study indicated that there may not have been an influence of L1 stress for all participants in the Indo-Aryan language group in their segmentation of the ALL materials. The Pashto and Urdu listeners performed poorly as expected, each answering just 29.17% of the trials correctly. However, the Hindi listener, who might have also been expected to perform poorly on the test, answered 70.83% of the trials correctly. In contrast, the Bengali listener, who might have been expected to benefit

---

17 However, a syllable’s nucleus is shortened when it loses its stress in morpho-phonological alternations. For example, baat = word, ba’taa = tell (Baart, 2014).

18 According to Chakraborty & Goffman (2011) there have been no studies on the phonetic realisation of stress in Bengali. Indeed, none were found by the author in her review of the relevant literature.
from the word initial pitch movement of both their first and second languages, 
answered just 33.33% of trials correctly.

The training received by the listeners does not appear to have influenced the 
outcome. Three of the Indo-Aryan participants received the training in connected 
speech processes. The Urdu and Bengali listeners did not make consistent use of the 
word initial pitch rise cue, though the Bengali listener might have been expected to 
do so. In contrast, the Hindi listener, who also received training in connected speech 
processes, performed well at the task, though word initial pitch rise is not a feature of 
their L1. Clearly, the training in connected speech processes did not have a consistent 
effect on the outcomes of these listeners. Only one Indo-Aryan participant received 
the word-spotting training. Word-spotting did not train the Pashto listener to use the 
word initial pitch rise cue.

Length of residence may provide an explanation for the good performance of the 
Hindi listener. With 19 months of residence in Australia, they had gained more 
exposure to English word initial pitch rise than the poor performing Pashto and Urdu 
listeners, who had resided in Australia for 12 months and 3 months, respectively. 
However, the poor performing Bengali listener had resided in Australia for 22 months 
and also had the benefit of exposure to both L1 and L2 word initial pitch rise. It must 
be noted, with even two native English listeners obtaining the low scores of 29.17% 
and 37.5% respectively, it is difficult to form conclusions based on the performance of 
one Bengali participant.
9.3.2 Word-Spotting Scores

9.3.2.1 Mean Percentage Correct

The mean percentage of correct word-spotting responses by the two language groups (L1 and L2) across the three daily training sets were analysed using a mixed repeated measures analysis of variance (ANOVA). As hypothesised, there was a significant main effect for 'language', $F(2, 23) = 95.61, p < 0.001$, with the L1 group receiving higher scores than the L2 group across all three training sets. The main effect of 'training set', and the 'training set' by 'language' interaction were not significant. Mean scores for each language group across the three training sets are shown in Figure 9.2.

As can be seen in Figure 9.2, the scores of the L2 group declined from Set A02 to Set A03. This result is similar to that of Experiment Four, in which there was a 4% fall in mean scores from Set A02 to Set A03, and is most likely due to differences in the difficulty of the training items and subsequent errors due to such factors as vowel
misperception. The L1 group did not experience the same difficulty with Set A03 and the mean scores of that group increased slightly across the three training sets.

### 9.3.2.2 Mean Percentage Segmented Correctly

A subsequent analysis was performed on the mean percentage of items segmented correctly. As shown in Experiment Four, analysing the data in this way removes the effects such as vowel and consonant misperception and focuses on the development of the targeted skill of speech segmentation.

This time the main effect of ‘training set’ was significant (shown here with a Greenhouse-Geisser adjustment to the degrees of freedom following a violation of Mauchly’s Test of Sphericity), $F(1.22, 26.82) = 5.64, p = .019$. Pairwise comparisons show that Set A03 was segmented significantly better than Set A01, $p = .024$.

![Figure 9.3 The mean percentage of items segmented correctly across the three training sets by the two language groups. The items in Set A03 were segmented significantly better than those in Set A01, $p = 0.24$. N=24](image)

The main effect of ‘language’ was also significant, $F(2, 23) = 54.01, p = <0.001$, with mean scores for the L1 group 10.44% higher than those of the L2 group. However, the interaction of ‘training set by language’ was not significant. The mean percentage of correctly segmented items for each group across the three sets is shown in Figure 9.3.
9.3.3 Correlations

There were no significant correlations between the participants’ ALL task scores and their ability to segment the word-spotting material.

9.4 Discussion

One possible explanation for the lack of a statistically significant influence of word-spotting training on ALL scores comes from a consideration of the different ways stress cues were realised in the training and post-test materials. As discussed previously in this chapter, in the word-spotting material, items in the Stress condition were recorded with increased pitch, duration and amplitude of stressed syllables and reduced vowels in unstressed syllables, whereas stress in the ALL materials was realised only as a left-edge pitch rise. If the participants attended to vowel reduction or, perhaps, duration as an important cue in segmenting the Stress condition items in the word-spotting material, pitch rise would not provide additional benefit in segmenting the ALL material, beyond that afforded to those who received training in connected speech processes.

There was very little difference between the performances of participants in the ALL task in each L1/L2 combined training group. However, although not a significant result, the L2 group that had received training in word-spotting received a numerically higher mean ALL task score than the L2 group that received training in connected speech processes. This further demonstrates that the word-spotting task holds promise as a method of training speech segmentation to learners of English.

The value of the word-spotting training is made particularly clear upon analysis of the responses to the task. Once more, analyses based on the participants’ segmentation of word-spotting items have indicated that segmentation skills improve over the
training sets. The steady improvements in segmentation made across the sets are distinct from overall performance. The L2 group’s mean score decreased from Set A02 to A03 in terms of the number of correct responses made, but there was an increase in the number of correct segmentations. Recall that the same pattern was seen in Experiment Four. This indicates that the word-spotting training is improving segmentation specifically, as it was designed to do.

Segmentation scores also improved for the L1 group. However, the L1 group’s mean score increased by only 1.75%, from Set A01 to Set A03, while the L2 group’s score improved by 5.92%. The improvement in the L2 groups’ segmentation scores despite a decrease in their correct responses and the disparity in the amount of improvement made by the two language groups suggest that mere familiarity with the task is not driving the significant difference between the training sets. These findings suggest that the segmentation skills of English learners improve across word-spotting training sets.
Chapter 10

General Discussion

10.1 Summary

The primary goal of this study was to determine if the word-spotting method could serve as a viable and beneficial method of training L2 speech segmentation skills. Speech segmentation is a fundamental process in listening comprehension, and inability to find word boundaries in a second language results in costs to speed and accuracy (Altenberg, 2005; Field, 2008c; Goh, 2000) and also contributes to learner anxiety (Bekleyen, 2009; Graham, 2006).

In Chapter 3, it was shown that a handful of writers have made a number of suggestions each on how to teach speech segmentation skills. However, with the exception of Al-jasser’s (2008) methodology, the efficacy of those suggestions have not been evaluated empirically. As English learners devote a substantial amount of time and money to improving their language skills, there is an ethical imperative to seek to establish the efficacy of teaching practices, and to focus on evidence-based practice where possible.

The study sought to establish if the word-spotting method would be more than or equally as effective as listening teaching methods that have been empirically shown to
improve L2 listening skills. This was assessed by measuring improvements in dictation scores, the number of missegmentations made in dictations, the reaction time to noticing newly taught words within continuous speech, and the ability to segment a continuous artificial language.

Problems were encountered in four of the five experiments presented in this thesis – ranging from issues with post-test choice and design to small sample-sizes. Although there were no significant differences between the results of experimental and control groups in this study, this project has provided good indications that the word-spotting method may provide training benefits and that further research is warranted to conclusively test its efficacy. The results also motivate recommendations about the most appropriate ways to test segmentation ability in the second language classroom.

The following sections of this chapter will provide an evaluation of the study, and conclude with a discussion on pedagogical implications for both training and testing, as well as recommendations for further research.

10.2 Evaluation

10.2.1 Viability

On the question of whether the word-spotting task is a viable method of training L2 speech segmentation, the results are clear. The validation of the task, reported in Chapter 4, showed that listeners from two rather different language backgrounds, Arabic and Mandarin, could perform the task well. All of the participants segmented at least 50% of the items correctly. Their results provided further evidence that L2 listeners are sensitive to the prosodic and phonotactic cues of English. Listeners had least difficulty with items with a weak-strong stress pattern in the first two syllables.
(for instance bicwrong and delfinger) and those containing an easy boundary transition (eastfid), and most difficulty with items with a strong-weak stress pattern (lipem and signaltev) and a difficult boundary transition (simkiss). Moreover, responses to the questionnaire showed that participants had a largely positive opinion of the task.

### 10.2.2 Efficacy

On the question of whether the word-spotting task is a beneficial method of training L2 speech segmentation, caution must be applied. As discussed above, small sample-sizes and issues with post-tests make interpretation of the results difficult. Due to the inability to adequately assess results in Experiments One and Two, tentative conclusions as to the efficacy of the word-spotting method will be drawn from the results of Experiments Three, Four and Five.

There was no significant difference between the results of the word-spotting training groups and the groups receiving training in connected speech processes on any measure. However, the significant improvement seen over time in Experiments Three and Four in increased dictation scores and fewer missegmentations for both groups is a positive sign. The word-spotting training was shown to be as effective as training in connected speech processes, the efficacy of which has been demonstrated in improving listening skills in English (Brown & Hilferty, 1986) and French (Kennedy & Blanchet, 2013). Therefore, by extension, the word-spotting training may also enhance listening skills. While the null hypothesis cannot be rejected, neither is it accepted.

An analysis of the responses to the word-spotting task itself showed that participants did improve in their performance with repeated exposure to the task. In Experiment
Three there was a very clear, statistically significant improvement in word-spotting scores over the six test sessions. In that experiment, the responses were coded only as correct or incorrect.

In Experiments Four and Five the word-spotting responses were also coded as to whether they were correctly segmented. This gave results beyond correct identification and orthographic representation of the word heard. Coding the results in this way removed aspects irrelevant to the question of segmentation, such as vowel perception, and thus increased the construct validity of the task in its use as a measure of speech segmentation skill.

The analyses based on the percentage of correct segmentations made in the word-spotting task established statistically that segmentation skills improve over the training sets, beyond the question of whether the item was identified correctly. Set A03 appeared to be more difficult than the other sets, with the percentage of correct responses falling from A02 to A03 in both Experiments Four and Five. However, the number of items segmented correctly increased from Set A02 to A03 in both experiments.

The L1 group did not experience the same difficulty with Set A03 in Experiment Five, with the mean number of correct responses increasing from Set A02 to Set A03. The disparity in performance between the two language groups suggests that Set A03 may contain a greater number of items with segmental features that pose difficulties for L2 listeners, leading to misperception. Still, despite the difficulty of the set, the L2 groups’ ability to segment the material improved.

It is important to note also that improvements can be seen within the first three sets, before the repetition of items commences in Set B01. This shows that the improvements reflect more than merely a gradual increase in familiarity with the task.
An effect of familiarity would benefit both groups equally, but the L2 group's segmentation scores increased more considerably across the three sets than the L1 group's scores. This was a numerical, rather than a statistically significant difference, but this result indicates that the word-spotting training improves speech segmentation as it has been designed to do. The word-spotting method shows promise as a method of training speech segmentation and improving second language listening skills.

10.2.3 Limitations of the Study

10.2.3.1 Low Participant Numbers

This study involved small classroom-based research, which is known to be problematic (Brown & Hilferty, 1986). The greatest hindrance in establishing the efficacy of the word-spotting training was the decline in participant numbers during the experiments. Ensuring commitment to the study was difficult. The classroom-based studies discussed in the review of the training literature in Chapter 3 incorporated the training into set coursework. In some instances, successful completion of the course was contingent on completion of the training. In contrast, participants in this study took part as an extracurricular activity at the end of each day, and were free to withdraw from the study at any time.

In every experiment but Experiment Five, they were paid just $5 an hour to commit to the training, or $2.50 a session. The participants initially found additional value in attending dedicated listening training, but as the demands of the College course naturally increased throughout the term, many participants felt they could no longer spare the time. This was understandable given that their entry to an Australian university depended on their final college grades.
10.2.3.2 Pre- and Post-tests

Different pre- and post-tests were used within experiments in order to avoid the confounding variable of a practice effect. However, potential differences in test difficulty in this study made the interpretation of results challenging. In retrospect, the same dictation sentences should have been used in the pre- and post-tests. The use of the control training group alongside the word-spotting experimental group would have allowed training effects to be separated from any practice effects.

10.2.3.3 Mini-Dictations

Questions remain about the reliability of the mini-dictations. The design was held as consistent as possible between tests over testing times, but the mean scores achieved by groups showed substantial variation from day to day.

Following Hughes (1989) participants were ranked according to their daily score. The participants’ rankings remained remarkably consistent across tests, with the most variation seen for the students with the lowest listening proficiency. This suggests that the tests may have been reliable internally, but differed in difficulty.

Ensuring tests contain the same number of sentences and that those sentences average the same number of words is clearly not enough to achieve uniformity between tests. It is likely the problem was caused by the total test length. The longer a well-balanced test is, the more reliable it will be (Hughes, 1989). With only ten sentences in each test and five in each dictation component, it is likely the mini-dictations were too short to provide adequate reliability.

There was a further problem caused by the inclusion of mini-dictations in the study. The decision was made to give daily mini-dictations in an effort to counter the negative effects of lack of attendance at post-tests. However, as dictation was a daily
feature of the program, its influence on post-test dictation improvements could not be ruled out. The testing device may well have become a training device.

10.2.3.4 Design of materials
The prevention of liaison occurring in the word-spotting materials (where word-final consonants were prevented from being articulated as the onset of following vowel-initial words) decreased the naturalness of the materials and thus, weakened the potential of the materials to provide training in segmenting words in accurate perceptual contexts. As noted earlier, such phonetic adjustments would result in acoustic-phonetic cues that could signal word boundaries, so the design of the materials in this way might have provided cues in the word-spotting training, but no assistance in segmenting these contexts in the dictation tasks. Any future experiments should assess the efficacy of materials featuring liaison across nonsense word/target boundaries.

10.4 Recommendations

10.4.1 Pedagogical Implications

10.4.1.1 Training
It is clear that L2 speech segmentation is a serious challenge for language learners and should receive more attention from second language listening researchers and teachers. The motivation to do so may be enhanced by the knowledge that L2 listeners are able to suppress L1 prosodic cues (Bradley et al., 1993; Hanuliková et al., 2011; Tremblay, et al., 2012) and to acquire L2 prosodic (Coughlin & Tremblay, 2012; Sanders et al., 2002; Suárez & Goh, 2013) and phonotactic cues (Al-jasser, 2008; Weber & Cutler, 2006). While only two studies have empirically tested methods of training L2 segmentation skills (this study and Al-jasser, 2008), further inspiration may come
from the numerous studies that have demonstrated the effectiveness of teaching decoding processes (including Al-jasser, 2008, Brown & Hilferty, 1986; Kennedy & Blanchet, 2013; Kiany & Shiramiry, 2002; Matthews & O’Toole, 2013).

Noting that the word-spotting task is one of the best tests of segmentation skills available, the significant improvements made by participants in segmenting the word-spotting material, itself, may encourage inclusion of the method in listening teaching sessions. As stated in Chapter 3, the word-spotting method could be incorporated into a program of small-scale teaching exercises with a decoding focus.

The task could be extended with a pre-listening stage of explicit instruction, perhaps an integration of Al-jasser’s (2008) activities to raise awareness of illicit English phonotactic sequences or one of the exercises Field (2008b) suggests to raise awareness of strong syllables as a boundary cue. A post-listening stage could draw listeners’ attention to the greater ease the class group likely had in segmenting strings with strong initial syllables and illicit phonotactic sequences, and encourage students to consciously draw upon this knowledge in the next word-spotting set or in a further segmentation exercise. As a later extension, the teacher could draw the class groups’ attention to the difficulty of placing word boundaries before weak syllables (the items where the weak nonsense syllable was in final position, as in lipem and signaltev) and use that as a springboard to discuss difficult segmentation contexts, such as words with prefixes. This could lead into an activity Field (2008b) suggests to train the recognition and segmentation of such words.

The order of presentation of the word-spotting items could be manipulated to gradually increase the difficulty of the task. Features to manipulate include the number of voices heard, boundary conditions included, and phonotactic sequences...
presented. The potential for this is discussed in greater detail in the section on recommendations for further research.

**Participants’ Reflections**

It must be noted that, in each experiment, participants preferred training in connected speech processes to the word-spotting method. However, recall that in the pilot test described in Chapter 4, in which participants received the word-spotting training individually rather than in a classroom setting, participants rated the training very highly.

As discussed previously, the preference for the training in connected speech processes may have been, in part, due to the training being conducted in a traditional teaching format. The use of the computer-based self-study format of the word-spotting method might make more sense to participants if training were conducted in their own time, or in a listening lab, rather than in the classroom.

Still, it must be acknowledged that the training in connected speech processes was very popular. Perhaps features of that training method could be incorporated into the pre-listening stage of a full listening teaching session, followed by word-spotting training in the while-listening stage.

**10.4.1.2 Testing**

As the research developed, a second focus emerged – one of designing methods of testing segmentation skills. Currently, the ALL and word-spotting tasks are considered the best ways to assess speech segmentation (Cutler, 2012). The word-spotting task has been used successfully as such a method for almost thirty years, but clearly a different task was required in this study. The ALL task can be used in a laboratory setting (as in Experiment Five), but it is not appropriate for the use of the teacher in the classroom.
Finding a suitable test of segmentation skills proved challenging. Important issues had to be considered. In particular, it was vital that the test have construct validity, in that it tested segmentation skills, and that successful completion of the test was achievable. Pedagogical implications of the tests of segmentation skills used and suggestions for classroom use follow.

**Listening Cloze Test**

The listening cloze test, adapted from Templeton (1977) and trialled in Experiment Two, was prohibitively difficult. However, a significant negative correlation between the number of partial dictation missegmentations made and the number of correct words in the listening cloze test indicated that accurate segmentation was indeed required to complete it. The test may have the potential for further development for classroom and research use with sentences of not more than seven words and a relaxed, but authentic speech rate. In order to gradually increase the complexity of the task, the tests could start with contextual sentences and progress to gradually more decontextualised sentences as students’ skills advance.

**Dictation**

The use of dictation in testing speech segmentation provides a way both to assess improvement and to diagnose segmentation problems. With lower-proficiency listeners, the use of partial dictations may have some benefits over whole dictations as some context may be given to make the task easier. In this study, training effects were more often found in analyses of partial dictation than whole dictation sentences.

When being used as a method for testing segmentation skills, dictations should not be assessed according to spelling or punctuation. If an assessment of segmentation skills is the only purpose in setting the dictation task, word-internal errors such as vowel substitutions should also be ignored.
Missegmentations

Counting the number of missegmentations made in dictations appears to be an effective way to test segmentation skills. The number of missegmentations fell in every post-test component but one, showing that participants were improving at segmenting the dictation input.

The negative correlations between missegmentation and dictation scores were significant in most instances, indicating that the more missegmentations were made, the fewer words were correctly recognised in dictations. This result was not wholly unexpected as the two scores are not independent. The correlations between partial dictation missegmentations and scores from the listening cloze test, and also between missegmentations made on day one of the mini-dictations and scores attained on day one of the word-spotting training, provide a more substantial indication that all three measures have validity as tests of segmentation ability.

Perhaps missegmentation analysis can have most promising potential as a diagnostic test. The location of difficulty in the dictation input and the type of missegmentation made in response often appeared to be predictable. This may be established through further research.

10.4.2 Recommendations for Further Research

The study provided indications that the word-spotting training may be an effective method of training second language listening skills. Any further study aiming to more conclusively test the method’s efficacy should address the potential for participant attrition as an important risk to the study’s viability. The study would be best conducted with a set classroom group, as a required component of listening instruction. Alternatively, the study could be conducted in the laboratory context common to most psycholinguistic experiments.
Another recommendation concerns the choice and use of post-tests. In classroom-based studies, the use of short partial dictation sentences is suggested. The same test sentences should be used in the pre- and post-tests, with control group results used to distinguish between training and practice effects.

In laboratory settings, the ALL task may be used. An important recommendation is that the ALL materials should be designed to reflect the word boundary cues trained in the word-spotting task. This means that initial syllables should not only contain a pitch rise, but also feature increased duration relative to other syllables. Boundary transitions should feature CC sequences in addition to CV sequences, and some of these should be illicit in English. The suggestion is clearly for an artificial language that more closely mirrors the structure of English. In order to test the relative contribution of word-spotting cues to improved segmentation skills, the prosodic and phonotactic cues could feature in two separate ALL languages.

There are several ways the insights gained from the initial study could be expanded upon. The material could be presented in sets of increasing difficulty. For instance, polysyllabic target words which form illegal phonotactic sequences beside the context (for example, traffic in trafficvoot) will be easier to detect, and thus could be presented earlier, than monosyllabic words with final or initial segments that form legal syllable internal sequences alongside the context (as in cost in enscost). Any increased benefit in structuring the training in this way could be measured. In addition, although no differences were found in the response to the material by Arabic and Mandarin listeners in the validation of the task, a future study could trial manipulation of the material for presentation to specific language groups.
Future research on the efficacy of the word-spotting task could take a different direction from that taken in this study. For instance, the training could be used to improve vowel perception. Many participants spoke about noticing a mismatch between vowels they had identified (and hence words they had typed in response), and the vowel in the input as revealed by the feedback given on the following screen. The repeated exposure, practice and feedback given in the word-spotting training could provide a useful way to fine tune vowel perception, and perhaps the perception of other segments and sequences.

This study made use of 513 words from an original list of 4372 words that did not contain phonetic embeddings. Many categories of words that were excluded from the present study due to complexity are available for use with native speakers, bilinguals or very advanced learners. Some of the words have already been recorded embedded in nonsense contexts in the Australian English accent. The full list of words forms a rich resource of words without any phonetic embeddings, for use in subsequent studies.

A further recommendation for future research concerns the output of the training and testing tasks in this study, which generated abundant data from listener responses. The analysis of this data could provide an enhanced understanding of L2 speech segmentation. Analyses may be conducted on the responses to the word-spotting training to ascertain whether there was any difference in the rate of improvement between the stress and phonotactic conditions. The data could also be investigated to establish whether participants of different L1 backgrounds segmented and improved upon the various phonotactic sequences differently. An analysis of the missegmentations generated in response to the dictation tasks may yield useful information about how second language learners perceive and segment English and
what input structures tend to result in the most difficulties for listeners. There were certain word sequences in the dictation input that created problems for many listeners and in some cases the same errors were made in response.

10. 5 Conclusion

This study documented an attempt to provide empirical evidence to support the use of an L2 listening teaching method. The word-spotting method for training speech segmentation was designed based on evidence from psycholinguistic research on the cues used by native English listeners to find word boundaries. As such, it is an example of the decoding approach to listening teaching, aiming to improve the use of L1 listening processes by L2 listeners in response to the structure of L1 speech.

The use of the word-spotting task as a method to train speech segmentation does show promise and can be recommended, with caution noted that further research is required to categorically establish its efficacy. No difference was found between control and experimental training groups at any point in this study, with significant improvements from pre- to post-test often found for both groups. This shows the word-spotting method is at least as beneficial as training in connected speech processes, which does have an empirical base. Perhaps more importantly, the significant improvements made by participants in segmenting the word-spotting material, beyond the selection of correct responses, provides considerable promise; especially as the word-spotting task is itself a well-respected test of segmentation skills.
References


Brown, J. D. (2006). Teacher resources for teaching connected speech. In J. D. Brown & K. Kondo-Brown (Eds.), *Perspectives on teaching connected speech to second language speakers* (pp. 27-50). Honolulu, HI: National Foreign Language Resource Center, University of Hawai‘i at Mānoa.


Rost, M. (2011). *Teaching and researching listening*. Harlow: Pearson Education.


248

# Words Excluded From Automated Search

## Part I

### Macquarie Dictionary IPA Codes

<table>
<thead>
<tr>
<th>IPA Symbol</th>
<th>Macquarie Code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Vowels</strong></td>
<td></td>
</tr>
<tr>
<td>i as in ‘peat’</td>
<td>i</td>
</tr>
<tr>
<td>i as in ‘pit’</td>
<td>1</td>
</tr>
<tr>
<td>e as in ‘pet’</td>
<td>8</td>
</tr>
<tr>
<td>æ as in ‘pat’</td>
<td>2</td>
</tr>
<tr>
<td>a as in ‘part’</td>
<td>a</td>
</tr>
<tr>
<td>ø as in ‘pot’</td>
<td>6</td>
</tr>
<tr>
<td>θ as in ‘but’</td>
<td>x</td>
</tr>
<tr>
<td>ð as in ‘port’</td>
<td>c</td>
</tr>
<tr>
<td>ø as in ‘put’</td>
<td>y</td>
</tr>
<tr>
<td>u as in ‘pool’</td>
<td>u</td>
</tr>
<tr>
<td>ə as in ‘pert’</td>
<td>3</td>
</tr>
<tr>
<td>ə as in ‘apart’</td>
<td>7</td>
</tr>
</tbody>
</table>
2. Consonants

θ as in ‘thin’ 0
ð as in ‘then’ q
ʃ as in ‘show’ 4
ʒ as in ‘measure’ 5
tʃ as in ‘choke’ t4
dʒ as in ‘joke’ d5
ŋ as in ‘sing’ 9
j as in ‘you’ j

3. Diphthongs

ai as in ‘buy’ a1
eɪ as in ‘bay’ e1
ɔɪ as in ‘boy’ c1
au as in ‘how’ ay
ɔʊ as in ‘hoe’ oy
ɪə as in ‘here’ 17
ɛə as in ‘hair’ 87
ʊə as in ‘tour’ y7
4. Foreign Sounds

<table>
<thead>
<tr>
<th>Sound</th>
<th>as in</th>
<th>French</th>
<th>Stress</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ó</td>
<td>as in</td>
<td>‘bon’</td>
<td>O</td>
</tr>
<tr>
<td>æ̃</td>
<td>as in</td>
<td>‘vin’</td>
<td>A</td>
</tr>
<tr>
<td>y</td>
<td>as in</td>
<td>‘rue’</td>
<td>!</td>
</tr>
</tbody>
</table>

5. Stress

- primary stress
- secondary stress
Part II

The following words were excluded from the automated search of words that contain embeddings of phonetic words. They are listed alongside the Macquarie Dictionary's own code versions of their IPA transcriptions. Please see Part I for details.

1. Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>IPA</th>
<th>Acronym</th>
<th>IPA</th>
<th>Acronym</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>AA</td>
<td>/e1 'e1/</td>
<td>DMus</td>
<td>/di 'mxs/</td>
<td>LAN</td>
<td>/l2n/</td>
</tr>
<tr>
<td>ADA</td>
<td>/e1 di e1/</td>
<td>DNS</td>
<td>/di 8n '8s/</td>
<td>LEM</td>
<td>/l8m/</td>
</tr>
<tr>
<td>AFI</td>
<td>/e1 8f 'a1/</td>
<td>DOS</td>
<td>/d6s/</td>
<td>LP</td>
<td>/8l 'pi/</td>
</tr>
<tr>
<td>AFL</td>
<td>/e1 8f '8l/</td>
<td>DSD</td>
<td>/di 8s 'di/</td>
<td>MAC</td>
<td>/m2k/</td>
</tr>
<tr>
<td>AHS</td>
<td>/e1 et4 '8s/</td>
<td>DSM</td>
<td>/di 8s '8m/</td>
<td>MD</td>
<td>/8m 'di/</td>
</tr>
<tr>
<td>AJ</td>
<td>/'e1 d5e1/</td>
<td>EFTA</td>
<td>/'8ft7/</td>
<td>mf.</td>
<td>/8m '8f/</td>
</tr>
<tr>
<td>AJA</td>
<td>/e1 d5ei 'e1/</td>
<td>EHF</td>
<td>/i et4 '8f/</td>
<td>MOS</td>
<td>/m6s/</td>
</tr>
<tr>
<td>AK</td>
<td>/e1 'ke1/</td>
<td>FJ</td>
<td>/8f'd5ei/</td>
<td>MX</td>
<td>/8m '8ks/</td>
</tr>
<tr>
<td>ALS</td>
<td>/e1 8l '8s/</td>
<td>FM</td>
<td>/8f'8m/</td>
<td>NAM</td>
<td>/8n e1 '8m/</td>
</tr>
<tr>
<td>AMF</td>
<td>/e1 8m '8f/</td>
<td>FX</td>
<td>/8f'8ks/</td>
<td>NB</td>
<td>/8n 'bi/</td>
</tr>
<tr>
<td>AO</td>
<td>/e1 'oy/</td>
<td>GATT</td>
<td>/g2t/</td>
<td>NBG</td>
<td>/8n bi 'd5i/</td>
</tr>
<tr>
<td>APRA</td>
<td>/'2pr7/</td>
<td>GEHA</td>
<td>/'giha/</td>
<td>NFF</td>
<td>/8n 8f '8f/</td>
</tr>
<tr>
<td>ASIO</td>
<td>/'eizioy/</td>
<td>G-G</td>
<td>/d5i-'d5i/</td>
<td>NGA</td>
<td>/8n d5i 'e1/</td>
</tr>
<tr>
<td>BAS</td>
<td>/b2z/</td>
<td>GIF</td>
<td>/gif/</td>
<td>NGO</td>
<td>/8n d5i 'oy/</td>
</tr>
<tr>
<td>CAMS</td>
<td>/k2mz/</td>
<td>GM</td>
<td>/d5i '8m/</td>
<td>NLF</td>
<td>/8n 8l '8f/</td>
</tr>
<tr>
<td>COAG</td>
<td>/'koy2g/</td>
<td>GNE</td>
<td>/d5i 8n 'i/</td>
<td>NMA</td>
<td>/8n 8m 'e1/</td>
</tr>
<tr>
<td>CYSS</td>
<td>/sis/</td>
<td>Guv</td>
<td>/gux/</td>
<td>NZ</td>
<td>/8n 'z8d/</td>
</tr>
<tr>
<td>DA</td>
<td>/di 'e1/</td>
<td>keV</td>
<td>/ke1 i 'vi/</td>
<td>OS</td>
<td>/oy '8s/</td>
</tr>
<tr>
<td>dd</td>
<td>/di 'di/</td>
<td>KG</td>
<td>/ke1 'd5i/</td>
<td>OAS</td>
<td>/oy e1 '8s/</td>
</tr>
<tr>
<td>DDH</td>
<td>/di di et4/</td>
<td>KL</td>
<td>/ke1 '8l/</td>
<td>OBE</td>
<td>/oy bi 'i/</td>
</tr>
<tr>
<td>DMR</td>
<td>/di 8m 'a/</td>
<td>KO</td>
<td>/ke1 'oy/</td>
<td>OE</td>
<td>/oy 'i/</td>
</tr>
</tbody>
</table>

253
### APPENDIX A

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>OED</td>
<td>/oy i ’di /</td>
</tr>
<tr>
<td>OEM</td>
<td>/oy i ’8m /</td>
</tr>
<tr>
<td>RADA</td>
<td>/’rad7 /</td>
</tr>
<tr>
<td>RAF</td>
<td>/ar e1 ’8f /</td>
</tr>
<tr>
<td>RBA</td>
<td>/a bi ’e1 /</td>
</tr>
<tr>
<td>RBL</td>
<td>/r b ’8l /</td>
</tr>
<tr>
<td>RD</td>
<td>/a ’di /</td>
</tr>
<tr>
<td>RDF</td>
<td>/a di ’8f /</td>
</tr>
<tr>
<td>RDO</td>
<td>/a di ’oy /</td>
</tr>
<tr>
<td>RGB</td>
<td>/a d5i ’bi /</td>
</tr>
<tr>
<td>RMO</td>
<td>/ar 8m ’oy /</td>
</tr>
<tr>
<td>RN</td>
<td>/ar ’8n /</td>
</tr>
<tr>
<td>RNA</td>
<td>/ar 8n ’e1 /</td>
</tr>
<tr>
<td>ROM</td>
<td>/ar 8n ’e1 /</td>
</tr>
<tr>
<td>RSL</td>
<td>/ar 8s ’8l /</td>
</tr>
<tr>
<td>RS</td>
<td>/ar ’8s /</td>
</tr>
<tr>
<td>RSJ</td>
<td>/ar 8s ’d5e1 /</td>
</tr>
<tr>
<td>RDF</td>
<td>/a di ’8f /</td>
</tr>
<tr>
<td>RSV</td>
<td>/ar 8s ’vi /</td>
</tr>
<tr>
<td>RV</td>
<td>/a ’vi /</td>
</tr>
<tr>
<td>RSM</td>
<td>/ar 8s ’8m /</td>
</tr>
<tr>
<td>RSV</td>
<td>/ar 8s ’vi /</td>
</tr>
<tr>
<td>RSV</td>
<td>/ar 8s ’vi /</td>
</tr>
<tr>
<td>RSV</td>
<td>/ar 8s ’vi /</td>
</tr>
<tr>
<td>RSV</td>
<td>/ar 8s ’vi /</td>
</tr>
<tr>
<td>SBS</td>
<td>/8s bi ’8s /</td>
</tr>
<tr>
<td>SDA</td>
<td>/8s di ’e1 /</td>
</tr>
<tr>
<td>SDA</td>
<td>/8s di ’e1 /</td>
</tr>
<tr>
<td>SKD</td>
<td>/8s ke1 ’di /</td>
</tr>
<tr>
<td>SLR</td>
<td>/8s 8l ’a /</td>
</tr>
<tr>
<td>SOG</td>
<td>/8s oy ’d5i /</td>
</tr>
<tr>
<td>SS</td>
<td>/8s ’8s /</td>
</tr>
<tr>
<td>VAD</td>
<td>/vi e1 ’di /</td>
</tr>
<tr>
<td>VD</td>
<td>/vi ’di /</td>
</tr>
<tr>
<td>VG</td>
<td>/vi ’d5i /</td>
</tr>
<tr>
<td>VMO</td>
<td>/vi 8m ’oy /</td>
</tr>
<tr>
<td>VNR</td>
<td>/vi 8n ’a /</td>
</tr>
<tr>
<td>VR</td>
<td>/vi ’a /</td>
</tr>
<tr>
<td>VS</td>
<td>/vi ’8s /</td>
</tr>
<tr>
<td>WAP</td>
<td>/w2p /</td>
</tr>
</tbody>
</table>

### 2. Letters

<table>
<thead>
<tr>
<th>Letter</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>/e1 /</td>
</tr>
<tr>
<td>A, a</td>
<td>/e1 /</td>
</tr>
<tr>
<td>D</td>
<td>/di /</td>
</tr>
<tr>
<td>D, d</td>
<td>/di /</td>
</tr>
<tr>
<td>E</td>
<td>/i /</td>
</tr>
<tr>
<td>E, e</td>
<td>/i /</td>
</tr>
<tr>
<td>f</td>
<td>/8f /</td>
</tr>
<tr>
<td>F, f</td>
<td>/8f /</td>
</tr>
<tr>
<td>G</td>
<td>/d5i /</td>
</tr>
<tr>
<td>G, g</td>
<td>/d5i /</td>
</tr>
<tr>
<td>H, h</td>
<td>/e1t4 /</td>
</tr>
<tr>
<td>J, j</td>
<td>/d5e1 /</td>
</tr>
<tr>
<td>K, k</td>
<td>/ke1 /</td>
</tr>
<tr>
<td>k</td>
<td>/ke1 /</td>
</tr>
<tr>
<td>L, l</td>
<td>/8l /</td>
</tr>
<tr>
<td>M</td>
<td>/8m /</td>
</tr>
<tr>
<td>M, m</td>
<td>/8m /</td>
</tr>
<tr>
<td>N, n</td>
<td>/8n /</td>
</tr>
<tr>
<td>O</td>
<td>/oy /</td>
</tr>
<tr>
<td>O, o</td>
<td>/oy /</td>
</tr>
<tr>
<td>R</td>
<td>/a /</td>
</tr>
<tr>
<td>R, r</td>
<td>/a /</td>
</tr>
<tr>
<td>S</td>
<td>/8s /</td>
</tr>
<tr>
<td>V</td>
<td>/vi /</td>
</tr>
<tr>
<td>V</td>
<td>/vi /</td>
</tr>
</tbody>
</table>

254
3. Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Pronunciation</th>
<th>Abbreviation</th>
<th>Pronunciation</th>
<th>Abbreviation</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abo</td>
<td>/ˈ2bɔɪ /</td>
<td>geo</td>
<td>/ˈdʒɔɪ /</td>
<td>prog</td>
<td>/ˈprɛɡ /</td>
</tr>
<tr>
<td>abs</td>
<td>/bɛz /</td>
<td>hep</td>
<td>/hɛp /</td>
<td>pud</td>
<td>/pjuːd /</td>
</tr>
<tr>
<td>alt</td>
<td>/ɔlt /</td>
<td>Jap</td>
<td>/dʒɛpt /</td>
<td>rad</td>
<td>/ræd /</td>
</tr>
<tr>
<td>app</td>
<td>/æp /</td>
<td>lav</td>
<td>/lɛv /</td>
<td>ref</td>
<td>/reɪf /</td>
</tr>
<tr>
<td>Aus</td>
<td>/ɔːz /</td>
<td>lib</td>
<td>/lɪb /</td>
<td>rep</td>
<td>/reɪp /</td>
</tr>
<tr>
<td>avo</td>
<td>/ˈəvoʊ /</td>
<td>Ls</td>
<td>/lɛz /</td>
<td>Reps</td>
<td>/rɛps /</td>
</tr>
<tr>
<td>biz</td>
<td>/bɪz /</td>
<td>mag</td>
<td>/mæɡ /</td>
<td>rev</td>
<td>/rɛv /</td>
</tr>
<tr>
<td>bod</td>
<td>/bɒd /</td>
<td>marg</td>
<td>/mærg /</td>
<td>REVs</td>
<td>/rɛvz /</td>
</tr>
<tr>
<td>bro</td>
<td>/broy /</td>
<td>meth</td>
<td>/mɛθ /</td>
<td>roid</td>
<td>/rɔɪd /</td>
</tr>
<tr>
<td>cam</td>
<td>/kæm /</td>
<td>mic</td>
<td>/maɪk /</td>
<td>roo</td>
<td>/rʊ /</td>
</tr>
<tr>
<td>camo</td>
<td>/ˈkæməʊ /</td>
<td>mod</td>
<td>/mɒd /</td>
<td>sav</td>
<td>/sæv /</td>
</tr>
<tr>
<td>comp</td>
<td>/kæmp /</td>
<td>Ms</td>
<td>/mɛz /</td>
<td>sec</td>
<td>/sɛk /</td>
</tr>
<tr>
<td>cuz</td>
<td>/kʌz /</td>
<td>Nat</td>
<td>/næt /</td>
<td>sep</td>
<td>/sɛp /</td>
</tr>
<tr>
<td>deb</td>
<td>/dɛb /</td>
<td>neg</td>
<td>/nɛɡ /</td>
<td>sept</td>
<td>/sɛpt /</td>
</tr>
<tr>
<td>Dis</td>
<td>/dɪs /</td>
<td>olˈ</td>
<td>/ɒl /</td>
<td>sis</td>
<td>/sɪs /</td>
</tr>
<tr>
<td>e.g.</td>
<td>/i ˈdʒi /</td>
<td>Oz</td>
<td>/ɔz /</td>
<td>spag</td>
<td>/spæɡ /</td>
</tr>
<tr>
<td>ed.</td>
<td>/ɛd /</td>
<td>pav</td>
<td>/pæv /</td>
<td>spiv</td>
<td>/spɪv /</td>
</tr>
<tr>
<td>exec</td>
<td>/ɛkˈsɛk /</td>
<td>pov</td>
<td>/pɔv /</td>
<td>sup</td>
<td>/sʌp /</td>
</tr>
<tr>
<td>fab</td>
<td>/fæb /</td>
<td>povo</td>
<td>/ˈpʌvəʊ /</td>
<td>tat</td>
<td>/tæt /</td>
</tr>
<tr>
<td>fave</td>
<td>/fɛv /</td>
<td>prem</td>
<td>/prɛm /</td>
<td>veg</td>
<td>/vɛɡ /</td>
</tr>
<tr>
<td>fess</td>
<td>/fɛs /</td>
<td>prep</td>
<td>/prɛp /</td>
<td>yer</td>
<td>/jɛr /</td>
</tr>
<tr>
<td>gen</td>
<td>/dʒɛn /</td>
<td>prof</td>
<td>/prɛf /</td>
<td>yˈr</td>
<td>/jɛr /</td>
</tr>
</tbody>
</table>
4. Exclamations

<table>
<thead>
<tr>
<th>Sound</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ah</td>
<td>/a/</td>
</tr>
<tr>
<td>aha</td>
<td>/a'ha/</td>
</tr>
<tr>
<td>ahem</td>
<td>/æ̆m/</td>
</tr>
<tr>
<td>bah</td>
<td>/ba/</td>
</tr>
<tr>
<td>bam</td>
<td>/bæm/</td>
</tr>
<tr>
<td>blah</td>
<td>/bla/</td>
</tr>
<tr>
<td>brrr</td>
<td>/br/</td>
</tr>
<tr>
<td>derr</td>
<td>/dɛ̈r/</td>
</tr>
<tr>
<td>duh</td>
<td>/dju/</td>
</tr>
<tr>
<td>eek</td>
<td>/i̯k/</td>
</tr>
<tr>
<td>eh</td>
<td>/e̯/</td>
</tr>
<tr>
<td>em</td>
<td>/ɛ̆m/</td>
</tr>
<tr>
<td>er</td>
<td>/ɛr/</td>
</tr>
<tr>
<td>ergh</td>
<td>/ɛrɡ/</td>
</tr>
<tr>
<td>gee</td>
<td>/ɡi/</td>
</tr>
<tr>
<td>geez</td>
<td>/ɡiz/</td>
</tr>
<tr>
<td>ha</td>
<td>/ha/</td>
</tr>
<tr>
<td>hah</td>
<td>/ha/</td>
</tr>
<tr>
<td>hmph</td>
<td>/hæmf/</td>
</tr>
<tr>
<td>huh</td>
<td>/hu/</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sound</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>jeez</td>
<td>/dʒɪz/</td>
</tr>
<tr>
<td>mmm</td>
<td>/m/</td>
</tr>
<tr>
<td>nah</td>
<td>/na/</td>
</tr>
<tr>
<td>oh</td>
<td>/ɔ̈/</td>
</tr>
<tr>
<td>oi</td>
<td>/oi/</td>
</tr>
<tr>
<td>ooh</td>
<td>/u/</td>
</tr>
<tr>
<td>oops</td>
<td>/ɔps/</td>
</tr>
<tr>
<td>ouch</td>
<td>/aʊtʃ/</td>
</tr>
<tr>
<td>pah</td>
<td>/pa/</td>
</tr>
<tr>
<td>pish</td>
<td>/paʃ/</td>
</tr>
<tr>
<td>psst</td>
<td>/pst/</td>
</tr>
<tr>
<td>rah</td>
<td>/ra/</td>
</tr>
<tr>
<td>sh</td>
<td>/ʃ/</td>
</tr>
<tr>
<td>tut</td>
<td>/tʌt/</td>
</tr>
<tr>
<td>ugh</td>
<td>/u/</td>
</tr>
<tr>
<td>uh</td>
<td>/uh/</td>
</tr>
<tr>
<td>um</td>
<td>/ʌm/</td>
</tr>
<tr>
<td>yah</td>
<td>/ja/</td>
</tr>
<tr>
<td>yep</td>
<td>/jɛp/</td>
</tr>
<tr>
<td>yup</td>
<td>/jʌp/</td>
</tr>
</tbody>
</table>
### 5. Function Words /Affixes

<table>
<thead>
<tr>
<th>Word</th>
<th>Affix</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>/æ/</td>
<td>as /æz/</td>
</tr>
<tr>
<td>ae</td>
<td>/æi/</td>
<td>be /bæ/</td>
</tr>
<tr>
<td>am</td>
<td>/æm/</td>
<td>-em /æm/</td>
</tr>
<tr>
<td>and</td>
<td>/ænd/</td>
<td>-er /ɔr/</td>
</tr>
<tr>
<td>are</td>
<td>/æ/</td>
<td>I /ai/</td>
</tr>
</tbody>
</table>

### 6. Infrequent Words

<table>
<thead>
<tr>
<th>Word</th>
<th>Affix</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>a'</td>
<td>/æ/</td>
<td>Ahab /æ-h2b/</td>
</tr>
<tr>
<td>aba</td>
<td>/'2b7/</td>
<td>ahoy /'h2c1/</td>
</tr>
<tr>
<td>abb</td>
<td>/'2b/</td>
<td>ail /'æl/</td>
</tr>
<tr>
<td>abba</td>
<td>/'2b7/</td>
<td>ain /'æn/</td>
</tr>
<tr>
<td>acca</td>
<td>/'2k7/</td>
<td>Ainu /'ænu/</td>
</tr>
<tr>
<td>Aceh</td>
<td>/'æ4e1/</td>
<td>aka /'ækæ/</td>
</tr>
<tr>
<td>acol</td>
<td>/'2k7l/</td>
<td>Akan /'akæn/</td>
</tr>
<tr>
<td>adat</td>
<td>/'x'dxt/</td>
<td>ala /'æl7/</td>
</tr>
<tr>
<td>ado</td>
<td>/'2d7u/</td>
<td>alb /'2lb/</td>
</tr>
<tr>
<td>ae</td>
<td>/'æi/</td>
<td>Alba /'2lb7/</td>
</tr>
<tr>
<td>aft</td>
<td>/'æft/</td>
<td>Alda /'2ld7/</td>
</tr>
<tr>
<td>afto</td>
<td>/'æftoy/</td>
<td>alf /'2lf/</td>
</tr>
<tr>
<td>aga</td>
<td>/'æg7/</td>
<td>alga /'ælg7/</td>
</tr>
<tr>
<td>agar</td>
<td>/'æg7/</td>
<td>Ali /'æli/</td>
</tr>
<tr>
<td>Agca</td>
<td>/'agk7/</td>
<td>alma /'2lm7/</td>
</tr>
<tr>
<td>agha</td>
<td>/'ag7/</td>
<td>aloo /'alu/</td>
</tr>
<tr>
<td>agin</td>
<td>/'g1n/</td>
<td>alu /'alu/</td>
</tr>
<tr>
<td>agog</td>
<td>/'g6g/</td>
<td>alum /'2l7m/</td>
</tr>
<tr>
<td>Agra</td>
<td>/'agr7/</td>
<td>Alva /'2lv7/</td>
</tr>
<tr>
<td>Arne</td>
<td>/an /</td>
<td>Bala</td>
</tr>
<tr>
<td>Arp</td>
<td>/ap /</td>
<td>Balt</td>
</tr>
<tr>
<td>aryl</td>
<td>/'2r1l /</td>
<td>bant</td>
</tr>
<tr>
<td>asp</td>
<td>/2sp /</td>
<td>bap</td>
</tr>
<tr>
<td>ass</td>
<td>/2s /</td>
<td>barf</td>
</tr>
<tr>
<td>Aube</td>
<td>/oyb /</td>
<td>Bax</td>
</tr>
<tr>
<td>auk</td>
<td>/ck /</td>
<td>baza</td>
</tr>
<tr>
<td>Auld</td>
<td>/cld /</td>
<td>beal</td>
</tr>
<tr>
<td>aum</td>
<td>/cm /</td>
<td>ben</td>
</tr>
<tr>
<td>Avar</td>
<td>/'ava /</td>
<td>berg</td>
</tr>
<tr>
<td>ave</td>
<td>/'ave1 /</td>
<td>berk</td>
</tr>
<tr>
<td>aver</td>
<td>/'v3 /</td>
<td>berm</td>
</tr>
<tr>
<td>Aves</td>
<td>/'e1viz /</td>
<td>Bes</td>
</tr>
<tr>
<td>avo</td>
<td>/'2voy /</td>
<td>Bhil</td>
</tr>
<tr>
<td>Avon</td>
<td>/'e1v7n /</td>
<td>biff</td>
</tr>
<tr>
<td>awa</td>
<td>/'wa /</td>
<td>bing</td>
</tr>
<tr>
<td>awl</td>
<td>/cl /</td>
<td>bip</td>
</tr>
<tr>
<td>awn</td>
<td>/cn /</td>
<td>bis</td>
</tr>
<tr>
<td>azo</td>
<td>/'e1zoy /</td>
<td>bisk</td>
</tr>
<tr>
<td>Azov</td>
<td>/'az6v /</td>
<td>bleb</td>
</tr>
<tr>
<td>baa</td>
<td>/ba /</td>
<td>blub</td>
</tr>
<tr>
<td>baal</td>
<td>/bal /</td>
<td>bob</td>
</tr>
<tr>
<td>baas</td>
<td>/bas /</td>
<td>Bock</td>
</tr>
<tr>
<td>bab</td>
<td>/b2b /</td>
<td>boco</td>
</tr>
<tr>
<td>baba</td>
<td>/'bab7 /</td>
<td>bode</td>
</tr>
<tr>
<td>baff</td>
<td>/b2f /</td>
<td>bole</td>
</tr>
<tr>
<td>baht</td>
<td>/bat /</td>
<td>boll</td>
</tr>
<tr>
<td>bal</td>
<td>/b2l /</td>
<td>Bon</td>
</tr>
</tbody>
</table>
Edom /'i dent / Eora /'itcr7 /
E'en /'in /
Eos /'i6s /
feme /f8m /
EEO /i i 'oy /
epos /'8p6s /
fen /f8n /
EEZ /i i 'z8d /
er /'7 /
Fens /f8nz /
eff /'8f /
-o y /
Fes /f8z /
eft /'8ft /
eger /'3g /
FET /f8t /
egad /i'g2d /
ergo /'3goy /
feta /'f8t7 /
eggis /'id57s /
Erin /'8ri /
fey /'fe1 /
eke /'ik /
Eris /'8ris /
fez /f8z /
Ekka /'8k7 /
erk /'3k /
rico /'fikoy /
Elba /'8lb7 /
erky /'3ki /
fid /f1d /
Elbe /'8lb /
err /'3 /
ifie /f1 /
elk /'8lk /
Erse /'3s /
fiief /'f1f /
ell /'8l /
erst /'3st /
fiife /f1af /
elm /'8lm /
ess /'8s /
fiisc /'f1sk /
Elul /'8'lul /
 ess /'8s /
Fisk /f1sk /
em /'8m /
eth /'8q /
flub /f1xb /
em /'7m /
etna /'8tn /
ofob /f6b /
-em /'7m /
Eure /'37 /
Foch /f64 /
emic /'im1k /
ex /'8ks /
fop /f6p /
Ems /'8mz /
eyas /'i7s /
Foy /f1c /
en /'8n /
Ezra /'8zr7 /
Fra /f1a /
Enga /'897 /
faff /'f2f /
Frau /fray /
Enid /'in7d /
fain /f1un /
fro /froy /
enol /'in6l /
falx /f2lks /
fug /f1xg /
Enos /'in6s /
fane /f1un /
Fuga /'fuga /
ens /'8nz /
farl /f1l /
fugu /'fugu /
EOKA /ei'oyk7 /
fay /'fe1 /
fut /f1x
futz /'fxts/  Gers /'gɛrs/  goss /'gɔs/  
Fyn /'fln/  geum /'dʒi7m/  Gov /'ɡʌv/  
gab /'ɡæb/  ghat /'ɡæt/  goy /'ɡɔi/  
Gaea /'dʒi7/  ghee /'ɡi/  Goya /'ɡɔi7/  
gaff /'ɡæf/  Gibb /'ɡɪb/  grad /'ɡræd/  
gaga /'ɡæɡə/  gid /'ɡɪd/  Graz /'ɡræz/  
Gaia /'ɡei7/  Gide /'ɡɪd/  grrl /'ɡɜrl/  
Gair /'ɡeə/  GIF /'ɡɪf/  Grus /'ɡrʌs/  
gal /'ɡæl/  Gifu /'ɡɪfʊ/  guan /'ɡwʌn/  
gam /'ɡæm/  ging /'ɡɪŋ/  gub /'ɡʌb/  
gan /'ɡæn/  gip /'dʒɪp/  guck /'ɡʌk/  
Gand /'ɡænd/  gird /'ɡɜrd/  guff /'ɡʌf/  
gar /'ɡær/  Gish /'ɡɪʃ/  Gui /'ɡuɪ/  
garb /'ɡærb/  gism /'dʒɪzm/  guk /'ɡʌk/  
Gard /'ɡɑrd/  gnat /'næt/  gunk /'ɡʌŋk/  
gare /'ɡær/  gnow /'ɡnɔ/  Gur /'ɡɜɾ/  
garn /'ɡɑrn/  gnu /'ɡnʌ/  Guv /'ɡʌv/  
Gary /'ɡeəri/  gob /'ɡəb/  gyp /'dʒɪp/  
gat /'ɡæt/  Gog /'ɡɔɡ/  hadj /'hædʒ/  
Gath /'ɡæθ/  Gogh /'ɡɔɡ/  haem /'heɪm/  
gaud /'ɡɔd/  gogs /'ɡɔɡs/  haft /'hæft/  
gaur /'ɡɔˈɑːr/  Gond /'ɡɔnd/  Hahn /'hæn/  
Gaya /'ɡeɪə/  gonk /'ɡɔŋk/  haj /'hæj/  
geld /'ɡeɪld/  goo /'ɡu/  haji /'hædʒi/  
gelt /'ɡeɪlt/  goof /'ɡʊf/  Han /'hæn/  
Genf /'ɡɛnʃ/  goog /'ɡʊɡ/  hank /'hæŋk/  
geo /'dʒiəʊ/  gook /'ɡʊk/  hap /'hæp/  
geri /'dʒiəri/  goom /'ɡʊm/  harl /'hærl/  

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>hath</td>
<td>/hɔt/</td>
</tr>
<tr>
<td>Idun</td>
<td>/ɪdʌn/</td>
</tr>
<tr>
<td>jis'</td>
<td>/dʒɪs/</td>
</tr>
<tr>
<td>heft</td>
<td>/hɛft/</td>
</tr>
<tr>
<td>Ifni</td>
<td>/ɪfni/</td>
</tr>
<tr>
<td>Jo</td>
<td>/dʒoʊ/</td>
</tr>
<tr>
<td>Hess</td>
<td>/hɛs/</td>
</tr>
<tr>
<td>ijmə</td>
<td>/ɪdʒma/</td>
</tr>
<tr>
<td>Joab</td>
<td>/dʒoʊb/</td>
</tr>
<tr>
<td>hest</td>
<td>/hɛst/</td>
</tr>
<tr>
<td>-im</td>
<td>/ɪm/</td>
</tr>
<tr>
<td>Joe</td>
<td>/dʒoʊ/</td>
</tr>
<tr>
<td>het</td>
<td>/hɛt/</td>
</tr>
<tr>
<td>Imo</td>
<td>/ɪmoʊ/</td>
</tr>
<tr>
<td>Joel</td>
<td>/dʒoʊl/</td>
</tr>
<tr>
<td>hex</td>
<td>/hɛks/</td>
</tr>
<tr>
<td>imp</td>
<td>/ɪmp/</td>
</tr>
<tr>
<td>joey</td>
<td>/dʒoʊj/</td>
</tr>
<tr>
<td>Hib</td>
<td>/hɪtə bi/</td>
</tr>
<tr>
<td>Ind</td>
<td>/ɪnd/</td>
</tr>
<tr>
<td>Jong</td>
<td>/dʒoʊŋ/</td>
</tr>
<tr>
<td>hic</td>
<td>/hɪk/</td>
</tr>
<tr>
<td>Inge</td>
<td>/ɪŋdʒ/</td>
</tr>
<tr>
<td>Jos</td>
<td>/dʒoʊs/</td>
</tr>
<tr>
<td>hob</td>
<td>/həb/</td>
</tr>
<tr>
<td>Ipoḥ</td>
<td>/ɪˈpoʊ/</td>
</tr>
<tr>
<td>Jose</td>
<td>/dʒoʊz/</td>
</tr>
<tr>
<td>hock</td>
<td>/hək/</td>
</tr>
<tr>
<td>Isar</td>
<td>/ɪˈzɑr/</td>
</tr>
<tr>
<td>joss</td>
<td>/dʒoʊs/</td>
</tr>
<tr>
<td>hod</td>
<td>/həd/</td>
</tr>
<tr>
<td>ism</td>
<td>/ɪzm/</td>
</tr>
<tr>
<td>Jung</td>
<td>/dʒʌŋ/</td>
</tr>
<tr>
<td>hog</td>
<td>/həɡ/</td>
</tr>
<tr>
<td>iwi</td>
<td>/ˈiwi/</td>
</tr>
<tr>
<td>Jura</td>
<td>/dʒʊrə/</td>
</tr>
<tr>
<td>Hogg</td>
<td>/hɔɡ/</td>
</tr>
<tr>
<td>iwis</td>
<td>/ˈiwiːs/</td>
</tr>
<tr>
<td>jus</td>
<td>/jʊs/</td>
</tr>
<tr>
<td>holf</td>
<td>/həlf/</td>
</tr>
<tr>
<td>Iwo</td>
<td>/ˈiwoʊ/</td>
</tr>
<tr>
<td>jusi</td>
<td>/ˈhʊsi/</td>
</tr>
<tr>
<td>hols</td>
<td>/həlz/</td>
</tr>
<tr>
<td>jag</td>
<td>/dʒæɡ/</td>
</tr>
<tr>
<td>kai</td>
<td>/keɪ/</td>
</tr>
<tr>
<td>Homs</td>
<td>/həms/</td>
</tr>
<tr>
<td>Jain</td>
<td>/dʒeɪn/</td>
</tr>
<tr>
<td>Kath</td>
<td>/kæθ/</td>
</tr>
<tr>
<td>hong</td>
<td>/hɔŋ/</td>
</tr>
<tr>
<td>Jane</td>
<td>/dʒeɪn/</td>
</tr>
<tr>
<td>Kay</td>
<td>/keɪ/</td>
</tr>
<tr>
<td>hoy</td>
<td>/hɔɪ/</td>
</tr>
<tr>
<td>jarl</td>
<td>/ˈɡaːl/</td>
</tr>
<tr>
<td>kaya</td>
<td>/ˈkɛjə/</td>
</tr>
<tr>
<td>hoya</td>
<td>/ˈhɔɪə/</td>
</tr>
<tr>
<td>Jat</td>
<td>/dʒæt/</td>
</tr>
<tr>
<td>kayo</td>
<td>/ˈkeɪoʊ/</td>
</tr>
<tr>
<td>hui</td>
<td>/ˈhui/</td>
</tr>
<tr>
<td>jay</td>
<td>/dʒeɪ/</td>
</tr>
<tr>
<td>keck</td>
<td>/kɛk/</td>
</tr>
<tr>
<td>Hui</td>
<td>/ˈhui/</td>
</tr>
<tr>
<td>Jeff</td>
<td>/dʒeɪf/</td>
</tr>
<tr>
<td>ked</td>
<td>/kɛd/</td>
</tr>
<tr>
<td>Hun</td>
<td>/hʌn/</td>
</tr>
<tr>
<td>Jeff</td>
<td>/dʒeɪf/</td>
</tr>
<tr>
<td>Jed</td>
<td>/kɛd/</td>
</tr>
<tr>
<td>hup</td>
<td>/hʌp/</td>
</tr>
<tr>
<td>Jena</td>
<td>/ˈdʒeɪnə/</td>
</tr>
<tr>
<td>Kef</td>
<td>/kɛf/</td>
</tr>
<tr>
<td>Hus</td>
<td>/hʌs/</td>
</tr>
<tr>
<td>Jens</td>
<td>/dʒeɪns/</td>
</tr>
<tr>
<td>Kemp</td>
<td>/kɛmp/</td>
</tr>
<tr>
<td>Huss</td>
<td>/hʌs/</td>
</tr>
<tr>
<td>Jes'</td>
<td>/dʒeɪs/</td>
</tr>
<tr>
<td>Ken</td>
<td>/kɛn/</td>
</tr>
<tr>
<td>ibu</td>
<td>/ˈɪbʊ/</td>
</tr>
<tr>
<td>Jess</td>
<td>/dʒeɪs/</td>
</tr>
<tr>
<td>Kent</td>
<td>/kɛnt/</td>
</tr>
<tr>
<td>idem</td>
<td>/ˈɪdɛm/</td>
</tr>
<tr>
<td>Jib</td>
<td>/dʒɪb/</td>
</tr>
<tr>
<td>Kerf</td>
<td>/kɛrf/</td>
</tr>
<tr>
<td>Idhi</td>
<td>/ˈɪdi/</td>
</tr>
<tr>
<td>Jim</td>
<td>/dʒɪm/</td>
</tr>
<tr>
<td>Kete</td>
<td>/ˈkɛti/</td>
</tr>
<tr>
<td>Ido</td>
<td>/ˈɪdəʊ/</td>
</tr>
<tr>
<td>Jip</td>
<td>/dʒɪp/</td>
</tr>
<tr>
<td>Khun</td>
<td>/kʰʌn/</td>
</tr>
</tbody>
</table>

262
<table>
<thead>
<tr>
<th>kike</th>
<th>kurg</th>
<th>lily</th>
<th>/kaɪk /</th>
<th>/ˈkɜrɡ /</th>
<th>/ˈlɪli /</th>
</tr>
</thead>
<tbody>
<tr>
<td>kine</td>
<td>kuri</td>
<td>Lima</td>
<td>/kaɪn /</td>
<td>/ˈkɜrɪ /</td>
<td>/ˈlɪmə /</td>
</tr>
<tr>
<td>kir</td>
<td>Kush</td>
<td>ling</td>
<td>/kɪɾ /</td>
<td>/ˈkʌʃ /</td>
<td>/lɪŋ /</td>
</tr>
<tr>
<td>kish</td>
<td>Kwa</td>
<td>loam</td>
<td>/kɪʃ /</td>
<td>/ˈkwɑ /</td>
<td>/ˈloʊm /</td>
</tr>
<tr>
<td>kith</td>
<td>la</td>
<td>Lod</td>
<td>/kɪθ /</td>
<td>/lɑ /</td>
<td>/ˈləd /</td>
</tr>
<tr>
<td>Klim</td>
<td>Laer</td>
<td>loll</td>
<td>/klɪm /</td>
<td>/ˈlɑːr /</td>
<td>/ˈlɒl /</td>
</tr>
<tr>
<td>knar</td>
<td>lah</td>
<td>loo</td>
<td>/nɑr /</td>
<td>/ˈlɑ /</td>
<td>/lu /</td>
</tr>
<tr>
<td>knop</td>
<td>lang</td>
<td>lop</td>
<td>/nɒp /</td>
<td>/ˈlɑŋ /</td>
<td>/lɒp /</td>
</tr>
<tr>
<td>knur</td>
<td>Lao</td>
<td>lope</td>
<td>/nɜ /</td>
<td>/ˈlɑʊ /</td>
<td>/ˈloʊp /</td>
</tr>
<tr>
<td>koan</td>
<td>Lara</td>
<td>Losh</td>
<td>/ˈkoʊn /</td>
<td>/ˈlɑrə /</td>
<td>/ˈlɒʃ /</td>
</tr>
<tr>
<td>Koch</td>
<td>lath</td>
<td>loup</td>
<td>/k64 /</td>
<td>/ˈlɑθ /</td>
<td>/lu /</td>
</tr>
<tr>
<td>koel</td>
<td>Laue</td>
<td>Loy</td>
<td>/ˈkoyl /</td>
<td>/ˈləʊə /</td>
<td>/ˈlɔɪ /</td>
</tr>
<tr>
<td>Kofu</td>
<td>lea</td>
<td>Loy</td>
<td>/ˈkɔfu /</td>
<td>/ˈliə /</td>
<td>/ˈlɔɪ /</td>
</tr>
<tr>
<td>kohal</td>
<td>Leah</td>
<td>Ls</td>
<td>/ˈkoʊhə /</td>
<td>/ˈliːə /</td>
<td>/ls /</td>
</tr>
<tr>
<td>konk</td>
<td>leat</td>
<td>luau</td>
<td>/k69k /</td>
<td>/ˈlɛt /</td>
<td>/ˈluɔ /</td>
</tr>
<tr>
<td>Koo</td>
<td>lech</td>
<td>luff</td>
<td>/ku /</td>
<td>/ˈlɛtʃ /</td>
<td>/lu /</td>
</tr>
<tr>
<td>Kos</td>
<td>lect</td>
<td>Luik</td>
<td>/k6 /</td>
<td>/ˈlɛkt /</td>
<td>/ˈluɪk /</td>
</tr>
<tr>
<td>kosh</td>
<td>lee</td>
<td>Luke</td>
<td>/k64 /</td>
<td>/li /</td>
<td>/ˈluk /</td>
</tr>
<tr>
<td>Kra</td>
<td>Lek</td>
<td>lulu</td>
<td>/kra /</td>
<td>/ˈlɛk /</td>
<td>/ˈluˌlu /</td>
</tr>
<tr>
<td>Krio</td>
<td>Leo</td>
<td>Luna</td>
<td>/ˈkriɔ /</td>
<td>/ˈlɛo /</td>
<td>/ˈluːnə /</td>
</tr>
<tr>
<td>kris</td>
<td>lerp</td>
<td>Lund</td>
<td>/kris /</td>
<td>/ˈlɜrp /</td>
<td>/ˈluːnd /</td>
</tr>
<tr>
<td>Krym</td>
<td>les</td>
<td>lune</td>
<td>/krɪm /</td>
<td>/ˈlɛs /</td>
<td>/ˈluːn /</td>
</tr>
<tr>
<td>kuia</td>
<td>leso</td>
<td>Lunt</td>
<td>/ˈkʊiə /</td>
<td>/ˈlɛso /</td>
<td>/ˈluːnt /</td>
</tr>
<tr>
<td>kuku</td>
<td>levy</td>
<td>Luo</td>
<td>/ˈkʊku /</td>
<td>/ˈlɛvi /</td>
<td>/ˈluo /</td>
</tr>
<tr>
<td>Kum</td>
<td>lex</td>
<td>luth</td>
<td>/kym /</td>
<td>/ˈlɛks /</td>
<td>/ˈluθ /</td>
</tr>
<tr>
<td>Kun</td>
<td>lezo</td>
<td>Lyon</td>
<td>/kun /</td>
<td>/ˈlɛzo /</td>
<td>/ˈluˌɔ /</td>
</tr>
<tr>
<td>kuna</td>
<td>Liao</td>
<td>Lys</td>
<td>/ˈkynə /</td>
<td>/ˈliəʊ /</td>
<td>/ˈlɪs /</td>
</tr>
<tr>
<td>Kura</td>
<td>lien</td>
<td>ma</td>
<td>/ˈkʊrə /</td>
<td>/ˈliən /</td>
<td>/ma /</td>
</tr>
</tbody>
</table>
maa /ma/  mirk /m3k/  murk /m3k/
Maas /mas/  mo /moy/  muss /mxs/
Mab /m2b/  mo /moy/  nab /n2b/
mac /m2k/  moa /moy/  nab /n2b/
mach /m2k/  moa /moy/  naff /'n2f/
Mack /m2k/  Moab /moy2b/  naga /'nag7/
Madl /madl/  Moe /moyi/  Naha /'naha/
mag /m2g/  moey /moyi/  Najd /n25d/
mal /m2l/  moi /mwa/  Nara /'nar7/
Mali /'mali/  moit /mcit/  nard /nad/
mana /'man7/  Moji /moy'd5i/  neb /n8b/
Manx /m29ks/  mole /moyk/  Nejd /n85d/
Mao /may/  moko /'moyko7/  nep /n8p/
mar /ma/  moll /m6l/  Ness /n8s/
Mara /'mar7/  molt /moylt/  nib /mb/
marl /mal/  mom /m6m/  nibs /n6b/
mass /m2s/  mon /m6n/  nig /nig/
Maui /'mayi/  mong /mx9/  nim /nim/
meg /m8g/  moo /mu/  Nin /nin/
mel /m8l/  moog /moyg/  Nith /n80/
meld /m8ld/  mot /moy/  nog /n6g/
melt /m8lt/  moue /mu/  nong /n69/
Meo /'me1oy/  mow /moy/  noo /nu/
mho /moy/  mowe /moy/  nosh /n64/
mick /m1k/  moxa /m6ks7/  nth /8no/
Ming /m19/  mozz /m6z/  nub /nxb/
mips /mips/  mu /mju/  Nuri /'ny7ri/
Mir /m17/  Muir /mij7/  o' /7/
<table>
<thead>
<tr>
<th>City</th>
<th>Pronunciation</th>
<th>City</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’</td>
<td>/øy /</td>
<td>op</td>
<td>/ɔp /</td>
</tr>
<tr>
<td>oast</td>
<td>/oyst /</td>
<td>opah</td>
<td>/’øyp7 /</td>
</tr>
<tr>
<td>Oban</td>
<td>/’øyb7n /</td>
<td>ope</td>
<td>/’øyp /</td>
</tr>
<tr>
<td>obi</td>
<td>/’øybi /</td>
<td>Ops</td>
<td>/’øyps /</td>
</tr>
<tr>
<td>Odra</td>
<td>/’cdra /</td>
<td>orc</td>
<td>/’cok /</td>
</tr>
<tr>
<td>o’er</td>
<td>/'c /</td>
<td>Ord</td>
<td>/’cokd /</td>
</tr>
<tr>
<td>ofay</td>
<td>/’øyfey /</td>
<td>ordo</td>
<td>/’øydoy /</td>
</tr>
<tr>
<td>oft</td>
<td>/’6ft /</td>
<td>ore</td>
<td>/’cok /</td>
</tr>
<tr>
<td>ogee</td>
<td>/’øyd5i /</td>
<td>Orel</td>
<td>/’6rəl /</td>
</tr>
<tr>
<td>Ogun</td>
<td>/’oygyn /</td>
<td>Orff</td>
<td>/’cəf /</td>
</tr>
<tr>
<td>ohm</td>
<td>/’oym /</td>
<td>Orne</td>
<td>/’cn /</td>
</tr>
<tr>
<td>oik</td>
<td>/’cik /</td>
<td>Oro</td>
<td>/’6roy /</td>
</tr>
<tr>
<td>Oita</td>
<td>/’øyit7 /</td>
<td>Orr</td>
<td>/’c /</td>
</tr>
<tr>
<td>Oka</td>
<td>/’oyk7 /</td>
<td>Orsk</td>
<td>/’esk /</td>
</tr>
<tr>
<td>okra</td>
<td>/’6kr7 /</td>
<td>ort</td>
<td>/’crt /</td>
</tr>
<tr>
<td>Okri</td>
<td>/’oykri /</td>
<td>oryx</td>
<td>/’6rɪks /</td>
</tr>
<tr>
<td>okta</td>
<td>/’6kt7 /</td>
<td>os</td>
<td>/’6s /</td>
</tr>
<tr>
<td>ol’</td>
<td>/’oyl /</td>
<td>Ossa</td>
<td>/’6s7 /</td>
</tr>
<tr>
<td>ole</td>
<td>/’oyl /</td>
<td>Otsu</td>
<td>/’cstu /</td>
</tr>
<tr>
<td>olio</td>
<td>/’oyli1y /</td>
<td>Oud</td>
<td>/’ayt /</td>
</tr>
<tr>
<td>olla</td>
<td>/’6l7 /</td>
<td>Oudh</td>
<td>/’ayd /</td>
</tr>
<tr>
<td>olm</td>
<td>/’oylm /</td>
<td>ouf</td>
<td>/’yf /</td>
</tr>
<tr>
<td>om</td>
<td>/’6m /</td>
<td>Oulu</td>
<td>/’aylu /</td>
</tr>
<tr>
<td>Omsk</td>
<td>/’6msk /</td>
<td>Ouse</td>
<td>/’uz /</td>
</tr>
<tr>
<td>oont</td>
<td>/’ynt /</td>
<td>oust</td>
<td>/’yast /</td>
</tr>
<tr>
<td>Oort</td>
<td>/’crt /</td>
<td>ouzo</td>
<td>/’uzəʊ /</td>
</tr>
<tr>
<td>Oost</td>
<td>/’oyst /</td>
<td>oy</td>
<td>/’ci /</td>
</tr>
<tr>
<td>oot</td>
<td>/’ut /</td>
<td>oyer</td>
<td>/’ciər /</td>
</tr>
<tr>
<td>plop</td>
<td>pl6p /</td>
<td>Ra</td>
<td>ra /</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>po</td>
<td>poy /</td>
<td>raff</td>
<td>r2f /</td>
</tr>
<tr>
<td>po</td>
<td>poy /</td>
<td>raga</td>
<td>'rag7 /</td>
</tr>
<tr>
<td>pock</td>
<td>p6k /</td>
<td>ragi</td>
<td>'r2gi /</td>
</tr>
<tr>
<td>poko</td>
<td>'poykoy /</td>
<td>raj</td>
<td>'rad5 /</td>
</tr>
<tr>
<td>Poe</td>
<td>poy /</td>
<td>raku</td>
<td>'raku /</td>
</tr>
<tr>
<td>poi</td>
<td>pc1 /</td>
<td>Rama</td>
<td>'ram7 /</td>
</tr>
<tr>
<td>poly</td>
<td>'p6li /</td>
<td>rata</td>
<td>'rat7 /</td>
</tr>
<tr>
<td>Pom</td>
<td>p6m /</td>
<td>Ree</td>
<td>ri /</td>
</tr>
<tr>
<td>poma</td>
<td>'p6m7 /</td>
<td>reft</td>
<td>r8ft /</td>
</tr>
<tr>
<td>pome</td>
<td>poym /</td>
<td>rels</td>
<td>r8lz /</td>
</tr>
<tr>
<td>pone</td>
<td>poyn /</td>
<td>rem</td>
<td>r8m /</td>
</tr>
<tr>
<td>pong</td>
<td>p69 /</td>
<td>reo</td>
<td>'rioy /</td>
</tr>
<tr>
<td>pouf</td>
<td>puf /</td>
<td>repp</td>
<td>r8p /</td>
</tr>
<tr>
<td>prad</td>
<td>pr2d /</td>
<td>ret</td>
<td>r8t /</td>
</tr>
<tr>
<td>Prag</td>
<td>prag /</td>
<td>rex</td>
<td>r8ks /</td>
</tr>
<tr>
<td>prau</td>
<td>pray /</td>
<td>Rey</td>
<td>re1 /</td>
</tr>
<tr>
<td>prow</td>
<td>pray /</td>
<td>rhea</td>
<td>ri7 /</td>
</tr>
<tr>
<td>Ptah</td>
<td>ta /</td>
<td>Rhee</td>
<td>ri /</td>
</tr>
<tr>
<td>puck</td>
<td>pxk /</td>
<td>Rhys</td>
<td>ris /</td>
</tr>
<tr>
<td>pud</td>
<td>pyd /</td>
<td>ria</td>
<td>ri7 /</td>
</tr>
<tr>
<td>pur</td>
<td>p3 /</td>
<td>riba</td>
<td>'rib7 /</td>
</tr>
<tr>
<td>purr</td>
<td>p3 /</td>
<td>rick</td>
<td>rik /</td>
</tr>
<tr>
<td>Pym</td>
<td>pum /</td>
<td>riel</td>
<td>r8l /</td>
</tr>
<tr>
<td>Qing</td>
<td>t419 /</td>
<td>Riga</td>
<td>'rig7 /</td>
</tr>
<tr>
<td>Qom</td>
<td>kym /</td>
<td>Rigi</td>
<td>'rigi /</td>
</tr>
<tr>
<td>qua</td>
<td>kwe1 /</td>
<td>rika</td>
<td>'r1k7 /</td>
</tr>
<tr>
<td>Qum</td>
<td>kym /</td>
<td>Rix</td>
<td>riks /</td>
</tr>
<tr>
<td>-------</td>
<td>---------------</td>
<td>-------</td>
<td>---------------</td>
</tr>
<tr>
<td>sala</td>
<td>/'salx/</td>
<td>skep</td>
<td>/sk8p/</td>
</tr>
<tr>
<td>SAM</td>
<td>/s2m/</td>
<td>skol</td>
<td>/sk6l/</td>
</tr>
<tr>
<td>san</td>
<td>/s2n/</td>
<td>skun</td>
<td>/skxn/</td>
</tr>
<tr>
<td>Sanz</td>
<td>/sanz/</td>
<td>slub</td>
<td>/slxb/</td>
</tr>
<tr>
<td>sard</td>
<td>/sad/</td>
<td>slue</td>
<td>/slu/</td>
</tr>
<tr>
<td>Saud</td>
<td>/sayd/</td>
<td>snib</td>
<td>/sn1b/</td>
</tr>
<tr>
<td>Sava</td>
<td>/'sav7/</td>
<td>snig</td>
<td>/sn1g/</td>
</tr>
<tr>
<td>saya</td>
<td>/'sajx/</td>
<td>snoz</td>
<td>/sn6z/</td>
</tr>
<tr>
<td>scag</td>
<td>/sk2g/</td>
<td>sny</td>
<td>/sna1/</td>
</tr>
<tr>
<td>sech</td>
<td>/s84/</td>
<td>sol</td>
<td>/s6l/</td>
</tr>
<tr>
<td>sen</td>
<td>/s8n/</td>
<td>Sol</td>
<td>/s6l/</td>
</tr>
<tr>
<td>Seth</td>
<td>/s8o/</td>
<td>sone</td>
<td>/soyn/</td>
</tr>
<tr>
<td>sett</td>
<td>/s8t/</td>
<td>sonk</td>
<td>/s69k/</td>
</tr>
<tr>
<td>shah</td>
<td>/4a/</td>
<td>sool</td>
<td>/sul/</td>
</tr>
<tr>
<td>Shan</td>
<td>/4an/</td>
<td>sop</td>
<td>/s6p/</td>
</tr>
<tr>
<td>shay</td>
<td>/4ei/</td>
<td>sot</td>
<td>/s6t/</td>
</tr>
<tr>
<td>Shem</td>
<td>/48m/</td>
<td>sou</td>
<td>/su/</td>
</tr>
<tr>
<td>shim</td>
<td>/4im/</td>
<td>sov</td>
<td>/s6v/</td>
</tr>
<tr>
<td>shiv</td>
<td>/41v/</td>
<td>stet</td>
<td>/st8t/</td>
</tr>
<tr>
<td>shmo</td>
<td>/4moy/</td>
<td>sui</td>
<td>/'sjui/</td>
</tr>
<tr>
<td>shog</td>
<td>/46g/</td>
<td>swy</td>
<td>/swa1/</td>
</tr>
<tr>
<td>Shri</td>
<td>/4ri/</td>
<td>tae</td>
<td>/te1/</td>
</tr>
<tr>
<td>Shub</td>
<td>/4yb/</td>
<td>taj</td>
<td>/tad5/</td>
</tr>
<tr>
<td>sib</td>
<td>/sib/</td>
<td>tam</td>
<td>/t2m/</td>
</tr>
<tr>
<td>Sim</td>
<td>/sim/</td>
<td>Tao</td>
<td>/day/</td>
</tr>
<tr>
<td>sith</td>
<td>/s10/</td>
<td>Tas</td>
<td>/t2z/</td>
</tr>
<tr>
<td>skag</td>
<td>/sk2g/</td>
<td>tau</td>
<td>/tay/</td>
</tr>
<tr>
<td>sked</td>
<td>/sk8d/</td>
<td>taw</td>
<td>/tc/</td>
</tr>
<tr>
<td>Word (Original)</td>
<td>Word (Phonetic)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ump /xmp /</td>
<td>vis /'v1z /</td>
<td>yeo /j0y /</td>
<td></td>
</tr>
<tr>
<td>umu /'umu /</td>
<td>von /f6n /</td>
<td>yez /j7z /</td>
<td></td>
</tr>
<tr>
<td>Ur /3 /</td>
<td>vox /v6ks /</td>
<td>Yid /j1d /</td>
<td></td>
</tr>
<tr>
<td>ur- /3 /</td>
<td>Wep /w8p /</td>
<td>Yin /j1n /</td>
<td></td>
</tr>
<tr>
<td>Uri /'uri /</td>
<td>wis /w1s /</td>
<td>yip /j1p /</td>
<td></td>
</tr>
<tr>
<td>ut /xt /</td>
<td>wop /w6p /</td>
<td>yo /j0y /</td>
<td></td>
</tr>
<tr>
<td>vac /v2k /</td>
<td>Wu /wu /</td>
<td>yoe /j0y /</td>
<td></td>
</tr>
<tr>
<td>vag /v2g /</td>
<td>xi /sa1 /</td>
<td>yon /j6n /</td>
<td></td>
</tr>
<tr>
<td>Var /va /</td>
<td>y' /j7 /</td>
<td>yow /jay /</td>
<td></td>
</tr>
<tr>
<td>vas /v2s /</td>
<td>ya /j7 /</td>
<td>Zea /'zi7 /</td>
<td></td>
</tr>
<tr>
<td>ve /v /</td>
<td>yap /j2p /</td>
<td>zo /z0y /</td>
<td></td>
</tr>
<tr>
<td>vee /vi /</td>
<td>yaw /jc /</td>
<td>zoa /'zoy7 /</td>
<td></td>
</tr>
<tr>
<td>vid /v1d /</td>
<td>ye /ji /</td>
<td>zot /z6t /</td>
<td></td>
</tr>
<tr>
<td>vim /v1m</td>
<td>yea /je1 /</td>
<td>Zug /tsug /</td>
<td></td>
</tr>
</tbody>
</table>
Full list of Embedded Word-Spotting Training Stimuli

Part I

Frequency levels 1 and 2

Initial Position

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>allow</td>
<td>allowplen</td>
<td>əˈlaʊplen</td>
</tr>
<tr>
<td>1</td>
<td>Asia</td>
<td>Asiadreen</td>
<td>əsɪəˈdrɪn</td>
</tr>
<tr>
<td>1</td>
<td>deep</td>
<td>deepralts</td>
<td>dɪpˈrælts</td>
</tr>
<tr>
<td>1</td>
<td>fell</td>
<td>fellprem</td>
<td>əˈfɛlprɛm</td>
</tr>
<tr>
<td>1</td>
<td>food</td>
<td>foodeeb</td>
<td>ˈfʊdɪb</td>
</tr>
<tr>
<td>1</td>
<td>foot</td>
<td>foot-thig</td>
<td>fʊtˈθɪɡ</td>
</tr>
<tr>
<td>1</td>
<td>foreign</td>
<td>foreigntrop</td>
<td>ˈfɔrɪŋtəp</td>
</tr>
<tr>
<td>1</td>
<td>free</td>
<td>freeblom</td>
<td>ˈfriːblɒm</td>
</tr>
<tr>
<td>1</td>
<td>gym</td>
<td>gympief</td>
<td>dʒɪmpɪf</td>
</tr>
<tr>
<td>1</td>
<td>hot</td>
<td>hotrel</td>
<td>hɒtˈrɛl</td>
</tr>
<tr>
<td>1</td>
<td>led</td>
<td>ledra</td>
<td>ˈlɛdɪərɛ</td>
</tr>
<tr>
<td>1</td>
<td>met</td>
<td>metsuf</td>
<td>mɛtʃʊf</td>
</tr>
<tr>
<td>1</td>
<td>ocean</td>
<td>oceandelbs</td>
<td>əʊˈʃændəlbz</td>
</tr>
<tr>
<td>1</td>
<td>out</td>
<td>outrel</td>
<td>əʊtˈrɛl</td>
</tr>
<tr>
<td>1</td>
<td>said</td>
<td>saidzim</td>
<td>ˈsɛdʒɪm</td>
</tr>
<tr>
<td>1</td>
<td>say</td>
<td>saybrom</td>
<td>ˈseɪbʁɔm</td>
</tr>
<tr>
<td>1</td>
<td>set</td>
<td>setwap</td>
<td>ˈsetwæp</td>
</tr>
<tr>
<td>1</td>
<td>shall</td>
<td>shalljoot</td>
<td>ʃældʒʊt</td>
</tr>
<tr>
<td>1</td>
<td>should</td>
<td>shouldthep</td>
<td>ʃʊdθɪp</td>
</tr>
<tr>
<td>1</td>
<td>step</td>
<td>steplut</td>
<td>ˈstɛplət</td>
</tr>
<tr>
<td>1</td>
<td>them</td>
<td>themfer</td>
<td>ˈðɛmφə</td>
</tr>
<tr>
<td>1</td>
<td>this</td>
<td>thiswup</td>
<td>ˈdɪswʊp</td>
</tr>
<tr>
<td>1</td>
<td>together</td>
<td>togethersnig</td>
<td>təˈgɛðərsnɪɡ</td>
</tr>
<tr>
<td>1</td>
<td>wing</td>
<td>wingthack</td>
<td>ˈwɪŋθæk</td>
</tr>
<tr>
<td>1</td>
<td>wool</td>
<td>woolstull</td>
<td>ˈwʊlstʊl</td>
</tr>
<tr>
<td>2</td>
<td>academic</td>
<td>academicreeg</td>
<td>æˈkəˈdemɪkɹɪɡ</td>
</tr>
<tr>
<td>2</td>
<td>alert</td>
<td>alertwid</td>
<td>əˈlɛtɹɪd</td>
</tr>
<tr>
<td>2</td>
<td>assembly</td>
<td>assemblybrazl</td>
<td>əˈsɛmˈbləɪbræzl</td>
</tr>
<tr>
<td>2</td>
<td>bent</td>
<td>bentain</td>
<td>ˈbɛntɛn</td>
</tr>
<tr>
<td>2</td>
<td>bet</td>
<td>betard</td>
<td>ˈbɛtərd</td>
</tr>
</tbody>
</table>
APPENDIX B

<table>
<thead>
<tr>
<th></th>
<th>Word</th>
<th>Phonetic</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>bury</td>
<td>bɜːriblɛŋk</td>
</tr>
<tr>
<td>1</td>
<td>cigarette</td>
<td>sigəˈrɛtrəv</td>
</tr>
<tr>
<td>1</td>
<td>mud</td>
<td>dɛtrɛnz</td>
</tr>
<tr>
<td>1</td>
<td>decent</td>
<td>disɛntɛm</td>
</tr>
<tr>
<td>1</td>
<td>deck</td>
<td>dɛklib</td>
</tr>
<tr>
<td>1</td>
<td>excess</td>
<td>əkˈseskwɔm</td>
</tr>
<tr>
<td>1</td>
<td>feather</td>
<td>fɛðəglaɛŋ</td>
</tr>
<tr>
<td>1</td>
<td>fiction</td>
<td>fɪkʃənθed</td>
</tr>
<tr>
<td>1</td>
<td>fog</td>
<td>fəɡok</td>
</tr>
<tr>
<td>1</td>
<td>frog</td>
<td>froɡnʃ</td>
</tr>
<tr>
<td>1</td>
<td>general</td>
<td>dʒɛnɛrklɪɡz</td>
</tr>
<tr>
<td>1</td>
<td>gang</td>
<td>ɡæŋkægz</td>
</tr>
<tr>
<td>1</td>
<td>hen</td>
<td>hɛnɛd</td>
</tr>
<tr>
<td>1</td>
<td>mess</td>
<td>mɛslʌnz</td>
</tr>
<tr>
<td>1</td>
<td>mix</td>
<td>mɪksɪʃ</td>
</tr>
<tr>
<td>1</td>
<td>moral</td>
<td>mɔrəlɪm</td>
</tr>
<tr>
<td>1</td>
<td>moss</td>
<td>mɔslɪɡ</td>
</tr>
<tr>
<td>1</td>
<td>objection</td>
<td>əbˈdʒekʃənθɔp</td>
</tr>
<tr>
<td>1</td>
<td>occur</td>
<td>əˈkɑːfrɛmz</td>
</tr>
<tr>
<td>1</td>
<td>owl</td>
<td>ɔʊlʃɪd</td>
</tr>
<tr>
<td>1</td>
<td>pack</td>
<td>pækɔm</td>
</tr>
<tr>
<td>1</td>
<td>profession</td>
<td>prəˈfɛʃəntən</td>
</tr>
<tr>
<td>1</td>
<td>reception</td>
<td>ˈrɛspəʃənup</td>
</tr>
<tr>
<td>1</td>
<td>regret</td>
<td>ˈrɛɡrɛtʌʊsp</td>
</tr>
<tr>
<td>1</td>
<td>shell</td>
<td>ʃɛlmən</td>
</tr>
<tr>
<td>1</td>
<td>star</td>
<td>stagrɪs</td>
</tr>
<tr>
<td>1</td>
<td>stir</td>
<td>stɪtl</td>
</tr>
<tr>
<td>1</td>
<td>suspicion</td>
<td>səˈspɪʃənɛld</td>
</tr>
</tbody>
</table>
### Easy Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>arm</td>
<td>armlek</td>
<td>amlek</td>
</tr>
<tr>
<td>1</td>
<td>death</td>
<td>deathmub</td>
<td>ɗθmab</td>
</tr>
<tr>
<td>1</td>
<td>desk</td>
<td>deskpoe</td>
<td>deskpoʊ</td>
</tr>
<tr>
<td>1</td>
<td>dish</td>
<td>dishdran</td>
<td>diʃdræn</td>
</tr>
<tr>
<td>1</td>
<td>distance</td>
<td>distanceveep</td>
<td>distænsvɪp</td>
</tr>
<tr>
<td>1</td>
<td>east</td>
<td>eastfɪd</td>
<td>ɪstfɪd</td>
</tr>
<tr>
<td>1</td>
<td>else</td>
<td>elsethet</td>
<td>ɛlʊðɛt</td>
</tr>
<tr>
<td>1</td>
<td>entrance</td>
<td>entrancethɪg</td>
<td>ɛntrænsθɪg</td>
</tr>
<tr>
<td>1</td>
<td>guess</td>
<td>guessruz</td>
<td>ɡɛsrʌz</td>
</tr>
<tr>
<td>1</td>
<td>hang</td>
<td>hangɛms</td>
<td>ʰæŋɛmz</td>
</tr>
<tr>
<td>1</td>
<td>happen</td>
<td>happenlur</td>
<td>ʰærpɛnɬ</td>
</tr>
<tr>
<td>1</td>
<td>head</td>
<td>headtra</td>
<td>ʰɛdtra</td>
</tr>
<tr>
<td>1</td>
<td>ink</td>
<td>inkfæp</td>
<td>ɪŋkɛp</td>
</tr>
<tr>
<td>1</td>
<td>job</td>
<td>jobclɔy</td>
<td>dʒɔbklɔi</td>
</tr>
<tr>
<td>1</td>
<td>laugh</td>
<td>laughjev</td>
<td>lɑfdʒɛv</td>
</tr>
<tr>
<td>1</td>
<td>on</td>
<td>onrɛl</td>
<td>ɔnrl</td>
</tr>
<tr>
<td>1</td>
<td>path</td>
<td>pathvɪg</td>
<td>paθvɪg</td>
</tr>
<tr>
<td>1</td>
<td>perhaps</td>
<td>perhapsthoun</td>
<td>ʰəˈhæpsθʌn</td>
</tr>
<tr>
<td>1</td>
<td>pig</td>
<td>pigkæŋ</td>
<td>ʰɪɡkæŋ</td>
</tr>
<tr>
<td>1</td>
<td>pound</td>
<td>poundfɪsp</td>
<td>.Serialize:1</td>
</tr>
<tr>
<td>1</td>
<td>rang</td>
<td>rangmɔʊl</td>
<td>ræŋmʊl</td>
</tr>
<tr>
<td>1</td>
<td>ship</td>
<td>shipgʊŋ</td>
<td>ʃɪpgʊŋ</td>
</tr>
<tr>
<td>1</td>
<td>sock</td>
<td>socknæl</td>
<td>ʃɔknæl</td>
</tr>
<tr>
<td>1</td>
<td>then</td>
<td>thenlɛz</td>
<td>ðɛnɬɛz</td>
</tr>
<tr>
<td>1</td>
<td>top</td>
<td>topgɛlp</td>
<td>топɡɛlp</td>
</tr>
<tr>
<td>1</td>
<td>traffic</td>
<td>trafficvoot</td>
<td>træfɪkʊt</td>
</tr>
<tr>
<td>1</td>
<td>wash</td>
<td>washpɪm</td>
<td>wɔʃpɪm</td>
</tr>
<tr>
<td>1</td>
<td>woman</td>
<td>womanrɛnz</td>
<td>wɔmnɛrnɬz</td>
</tr>
<tr>
<td>1</td>
<td>yet</td>
<td>jetɡʊk</td>
<td>ʃetɡʊk</td>
</tr>
<tr>
<td>2</td>
<td>aim</td>
<td>aɪmnoʊp</td>
<td>ɘimnɔp</td>
</tr>
<tr>
<td>2</td>
<td>analysis</td>
<td>anælæsɹvɛŋ</td>
<td>aˈnælæsɹvɛŋ</td>
</tr>
<tr>
<td>2</td>
<td>anxious</td>
<td>äŋʃæsɹvɪp</td>
<td>æŋʃæsɹvɪp</td>
</tr>
<tr>
<td>2</td>
<td>assemble</td>
<td>æsəmbɛrlɔlt</td>
<td>əˈsæmbɛrlɔlt</td>
</tr>
<tr>
<td>2</td>
<td>boss</td>
<td>bʊsθɛt</td>
<td>bʊsθɛt</td>
</tr>
<tr>
<td>2</td>
<td>cab</td>
<td>ˈkeɪbənz</td>
<td>ˈkeɪbənz</td>
</tr>
<tr>
<td>2</td>
<td>closet</td>
<td>closetvɪɡ</td>
<td>ˈkloʊzetvɪɡ</td>
</tr>
<tr>
<td>2</td>
<td>crop</td>
<td>krɒpɲɛʃ</td>
<td>krɒpɲɛʃ</td>
</tr>
<tr>
<td>2</td>
<td>deaf</td>
<td>defbɛz</td>
<td>defbɛz</td>
</tr>
<tr>
<td>2</td>
<td>dominant</td>
<td>dominæntmʌb</td>
<td>domænæntmʌb</td>
</tr>
<tr>
<td>2</td>
<td>dot</td>
<td>dotvɪp</td>
<td>ˈdɔtvɪp</td>
</tr>
<tr>
<td>2</td>
<td>ease</td>
<td>izplʌd</td>
<td>izplʌd</td>
</tr>
<tr>
<td>2</td>
<td>elect</td>
<td>aˈlɛktfɔp</td>
<td>aˈlɛktfɔp</td>
</tr>
<tr>
<td>2</td>
<td>elephant</td>
<td>elephantnen</td>
<td>ελεφαντνέν</td>
</tr>
<tr>
<td>---</td>
<td>----------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>2</td>
<td>exception</td>
<td>exceptionfets</td>
<td>ək'septʃənfets</td>
</tr>
<tr>
<td>2</td>
<td>exhibition</td>
<td>exhibitionris</td>
<td>ɛksə'bɪʃənris</td>
</tr>
<tr>
<td>2</td>
<td>experiment</td>
<td>experimentnom</td>
<td>ək'spərəməntnom</td>
</tr>
<tr>
<td>2</td>
<td>graph</td>
<td>graphthew</td>
<td>græfθu</td>
</tr>
<tr>
<td>2</td>
<td>guard</td>
<td>guardlubs</td>
<td>gadlʌbz</td>
</tr>
<tr>
<td>2</td>
<td>hook</td>
<td>hookchee</td>
<td>hoktʃi</td>
</tr>
<tr>
<td>2</td>
<td>hop</td>
<td>hopjuck</td>
<td>hɔpʤʌk</td>
</tr>
<tr>
<td>2</td>
<td>league</td>
<td>leaguevid</td>
<td>ligvid</td>
</tr>
<tr>
<td>2</td>
<td>lean</td>
<td>leanvave</td>
<td>lɪnveɪv</td>
</tr>
<tr>
<td>2</td>
<td>loss</td>
<td>lossthole</td>
<td>lɒstoʊl</td>
</tr>
<tr>
<td>2</td>
<td>oral</td>
<td>oralgarl</td>
<td>ərəlɡæl</td>
</tr>
<tr>
<td>2</td>
<td>possess</td>
<td>possessroch</td>
<td>pə'zɛsrətʃ</td>
</tr>
<tr>
<td>2</td>
<td>shed</td>
<td>shedlal</td>
<td>ʃedlæl</td>
</tr>
<tr>
<td>2</td>
<td>squirrel</td>
<td>squirrelthav</td>
<td>skwɪrəlðæv</td>
</tr>
<tr>
<td>2</td>
<td>target</td>
<td>targetgeen</td>
<td>tægætʃin</td>
</tr>
<tr>
<td>2</td>
<td>term</td>
<td>termlomb</td>
<td>ˈtæmlɒmb</td>
</tr>
<tr>
<td>2</td>
<td>thirst</td>
<td>thirstgri</td>
<td>θɛstɡri</td>
</tr>
<tr>
<td>2</td>
<td>wrist</td>
<td>wristklom</td>
<td>rɪstkʌm</td>
</tr>
<tr>
<td>2</td>
<td>yell</td>
<td>yellrem</td>
<td>ʃɛlrem</td>
</tr>
<tr>
<td>Frequency</td>
<td>Target Word</td>
<td>Training Word</td>
<td>IPA</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------</td>
<td>-----</td>
</tr>
<tr>
<td>1</td>
<td>art</td>
<td>artep</td>
<td>atəp</td>
</tr>
<tr>
<td>1</td>
<td>bed</td>
<td>bedesh</td>
<td>bɛdəʃ</td>
</tr>
<tr>
<td>1</td>
<td>drew</td>
<td>drewem</td>
<td>druəm</td>
</tr>
<tr>
<td>1</td>
<td>drop</td>
<td>dropel</td>
<td>dropəl</td>
</tr>
<tr>
<td>1</td>
<td>eat</td>
<td>eatej</td>
<td>itədʒ</td>
</tr>
<tr>
<td>1</td>
<td>enter</td>
<td>enterlesh</td>
<td>entəlæʃ</td>
</tr>
<tr>
<td>1</td>
<td>full</td>
<td>fullent</td>
<td>tʊlænt</td>
</tr>
<tr>
<td>1</td>
<td>got</td>
<td>gotesh</td>
<td>gotæʃ</td>
</tr>
<tr>
<td>1</td>
<td>hang</td>
<td>hangters</td>
<td>hæŋtəz</td>
</tr>
<tr>
<td>1</td>
<td>harbour</td>
<td>harbourjes</td>
<td>hæbədʒæs</td>
</tr>
<tr>
<td>1</td>
<td>hard</td>
<td>hardel</td>
<td>hædəl</td>
</tr>
<tr>
<td>1</td>
<td>have</td>
<td>havej</td>
<td>hævədʒ</td>
</tr>
<tr>
<td>1</td>
<td>her</td>
<td>hershel</td>
<td>hɛʃəl</td>
</tr>
<tr>
<td>1</td>
<td>lip</td>
<td>lipem</td>
<td>lɪpəm</td>
</tr>
<tr>
<td>1</td>
<td>map</td>
<td>mapef</td>
<td>mæpəf</td>
</tr>
<tr>
<td>1</td>
<td>open</td>
<td>openelt</td>
<td>ʊpənəlt</td>
</tr>
<tr>
<td>1</td>
<td>pot</td>
<td>potik</td>
<td>potik</td>
</tr>
<tr>
<td>1</td>
<td>present</td>
<td>presentesh</td>
<td>prəzəntæʃ</td>
</tr>
<tr>
<td>1</td>
<td>put</td>
<td>putrel</td>
<td>putrəl</td>
</tr>
<tr>
<td>1</td>
<td>room</td>
<td>roomel</td>
<td>rʊməl</td>
</tr>
<tr>
<td>1</td>
<td>salt</td>
<td>saltep</td>
<td>sɔltəp</td>
</tr>
<tr>
<td>1</td>
<td>second</td>
<td>secondef</td>
<td>sɛkəndəf</td>
</tr>
<tr>
<td>1</td>
<td>shook</td>
<td>shookens</td>
<td>sɔkənz</td>
</tr>
<tr>
<td>1</td>
<td>signal</td>
<td>signaltev</td>
<td>sɪgnəltəv</td>
</tr>
<tr>
<td>1</td>
<td>sing</td>
<td>singtent</td>
<td>sɪŋtənt</td>
</tr>
<tr>
<td>1</td>
<td>sir</td>
<td>sirding</td>
<td>sədɪŋ</td>
</tr>
<tr>
<td>1</td>
<td>spell</td>
<td>spellment</td>
<td>spɛlmənt</td>
</tr>
<tr>
<td>1</td>
<td>stood</td>
<td>stoodem</td>
<td>stʊdəm</td>
</tr>
<tr>
<td>1</td>
<td>sugar</td>
<td>sugarshen</td>
<td>ʃʊɡəʃən</td>
</tr>
<tr>
<td>1</td>
<td>sure</td>
<td>sureper</td>
<td>ʃʊrə</td>
</tr>
<tr>
<td>1</td>
<td>they</td>
<td>theyrep</td>
<td>ðeɪrəp</td>
</tr>
<tr>
<td>1</td>
<td>thick</td>
<td>thickrel</td>
<td>əɪkrəl</td>
</tr>
<tr>
<td>1</td>
<td>thing</td>
<td>thingent</td>
<td>θιν@Dataent</td>
</tr>
<tr>
<td>1</td>
<td>third</td>
<td>thirden</td>
<td>θάδαν</td>
</tr>
<tr>
<td>1</td>
<td>west</td>
<td>westeʃj</td>
<td>westædʒ</td>
</tr>
<tr>
<td>1</td>
<td>wood</td>
<td>woodel</td>
<td>woodel</td>
</tr>
<tr>
<td>2</td>
<td>absence</td>
<td>absencem</td>
<td>æbsænsəm</td>
</tr>
<tr>
<td>2</td>
<td>alcohol</td>
<td>alcoholent</td>
<td>ælkəholənt</td>
</tr>
<tr>
<td>2</td>
<td>bomb</td>
<td>bombesh</td>
<td>bɒməʃ</td>
</tr>
<tr>
<td>2</td>
<td>bother</td>
<td>botherel</td>
<td>bɒðərəl</td>
</tr>
<tr>
<td>2</td>
<td>bush</td>
<td>bushent</td>
<td>bɒʃənt</td>
</tr>
<tr>
<td>2</td>
<td>cigar</td>
<td>cigarel</td>
<td>sə'garəl</td>
</tr>
<tr>
<td>2</td>
<td>curse</td>
<td>cursef</td>
<td>kəsəf</td>
</tr>
<tr>
<td>2</td>
<td>depth</td>
<td>depthet</td>
<td>dɛpθət</td>
</tr>
<tr>
<td>2</td>
<td>dirt</td>
<td>dirtesh</td>
<td>dætəʃ</td>
</tr>
<tr>
<td>2</td>
<td>edge</td>
<td>edgeten</td>
<td>ɛdʒtən</td>
</tr>
<tr>
<td>2</td>
<td>emperor</td>
<td>emperortes</td>
<td>ɛmprərətəz</td>
</tr>
<tr>
<td>2</td>
<td>extra</td>
<td>extrashen</td>
<td>ɛkstrəʃən</td>
</tr>
<tr>
<td>2</td>
<td>fence</td>
<td>fencem</td>
<td>fənsəm</td>
</tr>
<tr>
<td>2</td>
<td>flesh</td>
<td>fleshters</td>
<td>fleʃtəz</td>
</tr>
<tr>
<td>2</td>
<td>jury</td>
<td>jurytet</td>
<td>dʒʊəritət</td>
</tr>
<tr>
<td>2</td>
<td>leisure</td>
<td>leisuregemp</td>
<td>lɛʒəgəmp</td>
</tr>
<tr>
<td>2</td>
<td>length</td>
<td>lengthesh</td>
<td>lɛŋθəʃ</td>
</tr>
<tr>
<td>2</td>
<td>measure</td>
<td>measureter</td>
<td>mεʒətə</td>
</tr>
<tr>
<td>2</td>
<td>medicine</td>
<td>medicinedred</td>
<td>mεdəsəndrid</td>
</tr>
<tr>
<td>2</td>
<td>melt</td>
<td>melttest</td>
<td>mɛltəst</td>
</tr>
<tr>
<td>2</td>
<td>mission</td>
<td>missionem</td>
<td>mɪʃənəm</td>
</tr>
<tr>
<td>2</td>
<td>object</td>
<td>objectel</td>
<td>ɔbdʒɛktəl</td>
</tr>
<tr>
<td>2</td>
<td>option</td>
<td>optionver</td>
<td>ɔpʃənəvə</td>
</tr>
<tr>
<td>2</td>
<td>owe</td>
<td>owement</td>
<td>oʊmənt</td>
</tr>
<tr>
<td>2</td>
<td>plow</td>
<td>plowmel</td>
<td>plɔʊməl</td>
</tr>
<tr>
<td>2</td>
<td>presence</td>
<td>presenceige</td>
<td>prezənsidʒ</td>
</tr>
<tr>
<td>2</td>
<td>rim</td>
<td>rimber</td>
<td>rimbə</td>
</tr>
<tr>
<td>2</td>
<td>stock</td>
<td>stockmer</td>
<td>stɔkme</td>
</tr>
<tr>
<td>2</td>
<td>tour</td>
<td>tourlem</td>
<td>tʊələm</td>
</tr>
<tr>
<td>2</td>
<td>treasure</td>
<td>treasuregend</td>
<td>trεʒərgənd</td>
</tr>
</tbody>
</table>
## Final Position

### Difficult Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>act</td>
<td>agact</td>
<td>ægækt</td>
</tr>
<tr>
<td>1</td>
<td>again</td>
<td>buldagain</td>
<td>βaldaˈɡɛn</td>
</tr>
<tr>
<td>1</td>
<td>agree</td>
<td>vamagree</td>
<td>vɪmæɡri</td>
</tr>
<tr>
<td>1</td>
<td>apple</td>
<td>vobapple</td>
<td>vɔbæpæl</td>
</tr>
<tr>
<td>1</td>
<td>cool</td>
<td>lencool</td>
<td>ɪɛŋkʊl</td>
</tr>
<tr>
<td>1</td>
<td>cost</td>
<td>enscost</td>
<td>ɛnskɔst</td>
</tr>
<tr>
<td>1</td>
<td>earth</td>
<td>rostearth</td>
<td>ˈrʊstæθ</td>
</tr>
<tr>
<td>1</td>
<td>English</td>
<td>bregenglish</td>
<td>bɛrɪŋɡɪlʃ</td>
</tr>
<tr>
<td>1</td>
<td>ill</td>
<td>vavill</td>
<td>ˈvævɪl</td>
</tr>
<tr>
<td>1</td>
<td>kept</td>
<td>feskept</td>
<td>fɛskɛpt</td>
</tr>
<tr>
<td>1</td>
<td>kick</td>
<td>merskick</td>
<td>mæskɪk</td>
</tr>
<tr>
<td>1</td>
<td>kiss</td>
<td>simkiss</td>
<td>sɪmkɪs</td>
</tr>
<tr>
<td>1</td>
<td>lent</td>
<td>tumplent</td>
<td>tæmplɛnt</td>
</tr>
<tr>
<td>1</td>
<td>long</td>
<td>slenklong</td>
<td>slɛŋklɔŋ</td>
</tr>
<tr>
<td>1</td>
<td>look</td>
<td>raslook</td>
<td>ræsloʊk</td>
</tr>
<tr>
<td>1</td>
<td>loud</td>
<td>pabloud</td>
<td>pæblaʊd</td>
</tr>
<tr>
<td>1</td>
<td>math</td>
<td>shasmath</td>
<td>ʃæsmaθ</td>
</tr>
<tr>
<td>1</td>
<td>men</td>
<td>vosmen</td>
<td>vɔsmɛn</td>
</tr>
<tr>
<td>1</td>
<td>most</td>
<td>oitsmost</td>
<td>ɔɪtsmɔst</td>
</tr>
<tr>
<td>1</td>
<td>noon</td>
<td>sesnoon</td>
<td>sɛsnʊn</td>
</tr>
<tr>
<td>1</td>
<td>oak</td>
<td>kalok</td>
<td>kælɔk</td>
</tr>
<tr>
<td>1</td>
<td>old</td>
<td>runold</td>
<td>rʊnɔld</td>
</tr>
<tr>
<td>1</td>
<td>pet</td>
<td>hespet</td>
<td>hɛspɛt</td>
</tr>
<tr>
<td>1</td>
<td>poem</td>
<td>rispoem</td>
<td>rɪspɔɪm</td>
</tr>
<tr>
<td>1</td>
<td>press</td>
<td>tispress</td>
<td>tɪspres</td>
</tr>
<tr>
<td>1</td>
<td>ring</td>
<td>detring</td>
<td>dɛtrɪŋ</td>
</tr>
<tr>
<td>1</td>
<td>red</td>
<td>geepred</td>
<td>gɪpred</td>
</tr>
<tr>
<td>1</td>
<td>tell</td>
<td>tastell</td>
<td>tæstɛl</td>
</tr>
<tr>
<td>1</td>
<td>test</td>
<td>nostest</td>
<td>nɒstɛst</td>
</tr>
<tr>
<td>1</td>
<td>tree</td>
<td>kestrree</td>
<td>kɛstrɪ</td>
</tr>
<tr>
<td>1</td>
<td>was</td>
<td>duswas</td>
<td>dʊswɔz</td>
</tr>
<tr>
<td>1</td>
<td>with</td>
<td>nentwith</td>
<td>nɛntwɪð</td>
</tr>
<tr>
<td>1</td>
<td>yes</td>
<td>fapyes</td>
<td>fæpɛjɛs</td>
</tr>
<tr>
<td>2</td>
<td>action</td>
<td>davaction</td>
<td>dævækʃən</td>
</tr>
<tr>
<td>2</td>
<td>anger</td>
<td>kelanger</td>
<td>kɛlæŋɡə</td>
</tr>
<tr>
<td>2</td>
<td>canal</td>
<td>fowscanal</td>
<td>fɔʊsˈkɑnl</td>
</tr>
<tr>
<td>2</td>
<td>comb</td>
<td>liscomb</td>
<td>lɪskɔm</td>
</tr>
<tr>
<td>2</td>
<td>complex</td>
<td>sescomplex</td>
<td>sɛskɔmplɛks</td>
</tr>
<tr>
<td>2</td>
<td>creation</td>
<td>biscreation</td>
<td>bɪskrɪˈɛʃən</td>
</tr>
<tr>
<td>2</td>
<td>eager</td>
<td>modeager</td>
<td>mɔdɪɡə</td>
</tr>
<tr>
<td>Element</td>
<td>Chameleon</td>
<td>ʃeəkələmənt</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>Tiserror</td>
<td>ʃisərə</td>
<td></td>
</tr>
<tr>
<td>Exist</td>
<td>Lifexist</td>
<td>ʃifəg'zɪst</td>
<td></td>
</tr>
<tr>
<td>Knee</td>
<td>Nisknee</td>
<td>ɲɪsni</td>
<td></td>
</tr>
<tr>
<td>Kettle</td>
<td>Joskettle</td>
<td>ʤɒpskɛtł</td>
<td></td>
</tr>
<tr>
<td>Lecture</td>
<td>Paglecture</td>
<td>ʰæglɛktfə</td>
<td></td>
</tr>
<tr>
<td>Neglect</td>
<td>Ropsneglect</td>
<td>ɾɒpsnɛ'glekt</td>
<td></td>
</tr>
<tr>
<td>Magnet</td>
<td>Bismagnet</td>
<td>bɪsmæɡnæt</td>
<td></td>
</tr>
<tr>
<td>Opera</td>
<td>Senopera</td>
<td>ʃɛnəprə</td>
<td></td>
</tr>
<tr>
<td>Opposite</td>
<td>Medopposite</td>
<td>mɛdəpəsæt</td>
<td></td>
</tr>
<tr>
<td>Ornament</td>
<td>Puvornament</td>
<td>ˈpəvənəmænt</td>
<td></td>
</tr>
<tr>
<td>Ounce</td>
<td>Gofounce</td>
<td>ɡəˈfɑuns</td>
<td></td>
</tr>
<tr>
<td>Pepper</td>
<td>Sispepper</td>
<td>ʃɪsɛpə</td>
<td></td>
</tr>
<tr>
<td>Poet</td>
<td>Hospoet</td>
<td>hɒspəʊət</td>
<td></td>
</tr>
<tr>
<td>Perfect</td>
<td>Fisperfect</td>
<td>ʃɪsəˈfɛkt</td>
<td></td>
</tr>
<tr>
<td>Pudding</td>
<td>Despudding</td>
<td>ˈdɛspədɪŋ</td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td>Respolish</td>
<td>ˈrɛspəlɪʃ</td>
<td></td>
</tr>
<tr>
<td>Resource</td>
<td>Deshresource</td>
<td>ˈdɛʃrəˈzɔs</td>
<td></td>
</tr>
<tr>
<td>Rhythm</td>
<td>Grofrhythm</td>
<td>ɡrofnəðəm</td>
<td></td>
</tr>
<tr>
<td>Tip</td>
<td>Sestip</td>
<td>ʃɛstɪp</td>
<td></td>
</tr>
<tr>
<td>Toss</td>
<td>Sestoss</td>
<td>ʃɛstəʊs</td>
<td></td>
</tr>
<tr>
<td>Whip</td>
<td>Ludwhip</td>
<td>ɭɔdwɪp</td>
<td></td>
</tr>
</tbody>
</table>
### Easy Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>back</td>
<td>motback</td>
<td>/mɒtˈbæk/</td>
</tr>
<tr>
<td>1</td>
<td>bag</td>
<td>gakbag</td>
<td>/ɡækˈbæɡ/</td>
</tr>
<tr>
<td>1</td>
<td>best</td>
<td>ludbest</td>
<td>/lədbɛst/</td>
</tr>
<tr>
<td>1</td>
<td>book</td>
<td>fotbook</td>
<td>/fɔtˈbɒk/</td>
</tr>
<tr>
<td>1</td>
<td>cherry</td>
<td>bapcherry</td>
<td>/bæpˈtʃɛri/</td>
</tr>
<tr>
<td>1</td>
<td>cook</td>
<td>birtcook</td>
<td>/bɜtˈkɒk/</td>
</tr>
<tr>
<td>1</td>
<td>could</td>
<td>seftcould</td>
<td>/sɛftˈkɔd/</td>
</tr>
<tr>
<td>1</td>
<td>cloth</td>
<td>detcloth</td>
<td>/dɛtˈklɔθ/</td>
</tr>
<tr>
<td>1</td>
<td>dead</td>
<td>tepdead</td>
<td>/tɛpˈdɛd/</td>
</tr>
<tr>
<td>1</td>
<td>dog</td>
<td>fubdog</td>
<td>/fʌbˈdɒɡ/</td>
</tr>
<tr>
<td>1</td>
<td>dress</td>
<td>hekdress</td>
<td>/hɛkˈdrɛs/</td>
</tr>
<tr>
<td>1</td>
<td>fix</td>
<td>glidfix</td>
<td>/ɡlɪdˈfɪks/</td>
</tr>
<tr>
<td>1</td>
<td>gather</td>
<td>smepgather</td>
<td>/sɪmɛpˈɡæðə/</td>
</tr>
<tr>
<td>1</td>
<td>get</td>
<td>rertget</td>
<td>/rɛrtˈɡɛt/</td>
</tr>
<tr>
<td>1</td>
<td>go</td>
<td>preetgo</td>
<td>/prɛtˈɡoʊ/</td>
</tr>
<tr>
<td>1</td>
<td>good</td>
<td>sibgood</td>
<td>/sɪbˈɡʊd/</td>
</tr>
<tr>
<td>1</td>
<td>grew</td>
<td>ketgrew</td>
<td>/kɛtˈɡrɛʊ/</td>
</tr>
<tr>
<td>1</td>
<td>may</td>
<td>talfmay</td>
<td>/tɛlfˈmeɪ/</td>
</tr>
<tr>
<td>1</td>
<td>miss</td>
<td>hetmiss</td>
<td>/hɛtmɪs/</td>
</tr>
<tr>
<td>1</td>
<td>moment</td>
<td>chotmoment</td>
<td>/tʃɔtˈmoʊmənt/</td>
</tr>
<tr>
<td>1</td>
<td>moon</td>
<td>thidmoon</td>
<td>/θɪdˈmʊn/</td>
</tr>
<tr>
<td>1</td>
<td>now</td>
<td>dutnow</td>
<td>/daʊtnəʊ/</td>
</tr>
<tr>
<td>1</td>
<td>nurse</td>
<td>bimnurse</td>
<td>/bɪmˈnɜrs/</td>
</tr>
<tr>
<td>1</td>
<td>pardon</td>
<td>deegpardon</td>
<td>/dɛɡˈpadn/</td>
</tr>
<tr>
<td>1</td>
<td>pay</td>
<td>nistpay</td>
<td>/nɪstˈpeɪ/</td>
</tr>
<tr>
<td>1</td>
<td>pull</td>
<td>ketpull</td>
<td>/kɛtˈpʊl/</td>
</tr>
<tr>
<td>1</td>
<td>three</td>
<td>divthree</td>
<td>/dɪvˈθri/</td>
</tr>
<tr>
<td>1</td>
<td>took</td>
<td>smagtook</td>
<td>/smaˈɡtɒk/</td>
</tr>
<tr>
<td>2</td>
<td>chemist</td>
<td>ditchemist</td>
<td>/dɪtˈkɛmɪst/</td>
</tr>
<tr>
<td>2</td>
<td>comma</td>
<td>eabcomma</td>
<td>/ɪbˈkɒmə/</td>
</tr>
<tr>
<td>2</td>
<td>congress</td>
<td>fepcongress</td>
<td>/fɛpˈkɔŋɡrɛs/</td>
</tr>
<tr>
<td>2</td>
<td>cushion</td>
<td>afcushion</td>
<td>/æfˈkʊʃən/</td>
</tr>
<tr>
<td>2</td>
<td>deed</td>
<td>hibdeed</td>
<td>/hɪbˈdɪd/</td>
</tr>
<tr>
<td>2</td>
<td>desert (noun)</td>
<td>bolkdesert</td>
<td>/bɔlˌkɛzət/</td>
</tr>
<tr>
<td>2</td>
<td>desperate</td>
<td>fæpdesperate</td>
<td>/fæpˈdɛspræt/</td>
</tr>
<tr>
<td>2</td>
<td>dessert</td>
<td>lutdessert</td>
<td>/lʊtˌdɛzət/</td>
</tr>
<tr>
<td>2</td>
<td>dig</td>
<td>bipdig</td>
<td>/bɪpˈdɪɡ/</td>
</tr>
<tr>
<td>2</td>
<td>fist</td>
<td>higfist</td>
<td>/hɪɡˈfɪst/</td>
</tr>
<tr>
<td>2</td>
<td>frost</td>
<td>magfrost</td>
<td>/mæɡˈfrost/</td>
</tr>
<tr>
<td>2</td>
<td>gap</td>
<td>wushgap</td>
<td>/wʌʃˈɡæp/</td>
</tr>
<tr>
<td>2</td>
<td>garage</td>
<td>pemgarage</td>
<td>/pɛmˈɡærə/</td>
</tr>
<tr>
<td>2</td>
<td>golf</td>
<td>flutgolf</td>
<td>flʌtɡɔlf</td>
</tr>
<tr>
<td>2</td>
<td>goose</td>
<td>sotgoose</td>
<td>sɒtɡʊs</td>
</tr>
<tr>
<td>2</td>
<td>grab</td>
<td>geetgrab</td>
<td>ɡɪtɡræb</td>
</tr>
<tr>
<td>2</td>
<td>jar</td>
<td>keefjar</td>
<td>kɪfdʒə</td>
</tr>
<tr>
<td>2</td>
<td>jaw</td>
<td>snopjaw</td>
<td>snɒpʤɔ</td>
</tr>
<tr>
<td>2</td>
<td>leap</td>
<td>lenleap</td>
<td>ɪɛnliŋ</td>
</tr>
<tr>
<td>2</td>
<td>memory</td>
<td>gopmemory</td>
<td>ɡɒpˈmɛməri</td>
</tr>
<tr>
<td>2</td>
<td>mood</td>
<td>eentmood</td>
<td>ɪntmud</td>
</tr>
<tr>
<td>2</td>
<td>myth</td>
<td>ligmyth</td>
<td>lɪgmiθ</td>
</tr>
<tr>
<td>2</td>
<td>noodle</td>
<td>lemnoodele</td>
<td>ɪɛmnuːdəl</td>
</tr>
<tr>
<td>2</td>
<td>noun</td>
<td>motnoun</td>
<td>mɒtnaʊn</td>
</tr>
<tr>
<td>2</td>
<td>pop</td>
<td>jootpop</td>
<td>dʒuːtpɔp</td>
</tr>
<tr>
<td>2</td>
<td>position</td>
<td>ockposition</td>
<td>ɒkˈpəzɪʃən</td>
</tr>
<tr>
<td>2</td>
<td>powder</td>
<td>dutpowder</td>
<td>dʌtpəʊdə</td>
</tr>
<tr>
<td>2</td>
<td>rank</td>
<td>vormrank</td>
<td>vɔrˈmæŋk</td>
</tr>
<tr>
<td>2</td>
<td>reference</td>
<td>hunreference</td>
<td>hʌnˈrefrəns</td>
</tr>
<tr>
<td>2</td>
<td>resist</td>
<td>senresist</td>
<td>sənˈresɪst</td>
</tr>
<tr>
<td>2</td>
<td>tap</td>
<td>gerptap</td>
<td>ɡærptæp</td>
</tr>
<tr>
<td>2</td>
<td>theme</td>
<td>elttheme</td>
<td>ɛltθɪm</td>
</tr>
</tbody>
</table>
## Stress Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>absent</td>
<td>lepabsent</td>
<td>/ləpˈæbsənt/</td>
</tr>
<tr>
<td>1</td>
<td>angry</td>
<td>simangry</td>
<td>/sɪmˈæŋɡri/</td>
</tr>
<tr>
<td>1</td>
<td>big</td>
<td>kebig</td>
<td>/keˈbɪɡ/</td>
</tr>
<tr>
<td>1</td>
<td>day</td>
<td>piday</td>
<td>/pɪˈdeɪ/</td>
</tr>
<tr>
<td>1</td>
<td>ear</td>
<td>jikear</td>
<td>/dʒɪkˈɪə/</td>
</tr>
<tr>
<td>1</td>
<td>end</td>
<td>shenend</td>
<td>/ʃənˈɛnd/</td>
</tr>
<tr>
<td>1</td>
<td>finger</td>
<td>delfinger</td>
<td>/dɛlˈfɪŋɡə/</td>
</tr>
<tr>
<td>1</td>
<td>habit</td>
<td>kehabit</td>
<td>/kɛlˈhæbət/</td>
</tr>
<tr>
<td>1</td>
<td>hard</td>
<td>seshard</td>
<td>/səsˈhɑd/</td>
</tr>
<tr>
<td>1</td>
<td>half</td>
<td>tivhalf</td>
<td>/tɪvˈhɑf/</td>
</tr>
<tr>
<td>1</td>
<td>if</td>
<td>chesif</td>
<td>/tʃəzˈɪf/</td>
</tr>
<tr>
<td>1</td>
<td>king</td>
<td>treking</td>
<td>/tɾeˈkɪŋ/</td>
</tr>
<tr>
<td>1</td>
<td>neck</td>
<td>treneck</td>
<td>/tɾeˈnɛk/</td>
</tr>
<tr>
<td>1</td>
<td>nest</td>
<td>menest</td>
<td>/məˈnɛst/</td>
</tr>
<tr>
<td>1</td>
<td>off</td>
<td>cheboff</td>
<td>/tʃəˈbɒf/</td>
</tr>
<tr>
<td>1</td>
<td>push</td>
<td>repush</td>
<td>/rɛˈpʊʃ/</td>
</tr>
<tr>
<td>1</td>
<td>rather</td>
<td>fidrather</td>
<td>/fɪdˈrɑθ/</td>
</tr>
<tr>
<td>1</td>
<td>rest</td>
<td>delrest</td>
<td>/dəlˈrɛst/</td>
</tr>
<tr>
<td>1</td>
<td>roof</td>
<td>wijroof</td>
<td>/wɪdʒˈrʊf/</td>
</tr>
<tr>
<td>1</td>
<td>sang</td>
<td>chersang</td>
<td>/tʃəˈsæŋ/</td>
</tr>
<tr>
<td>1</td>
<td>shock</td>
<td>relshock</td>
<td>/rɛlˈʃɒk/</td>
</tr>
<tr>
<td>1</td>
<td>shop</td>
<td>melshop</td>
<td>/mɛlˈʃɒp/</td>
</tr>
<tr>
<td>1</td>
<td>smell</td>
<td>netsmell</td>
<td>/nɛtˈsmɛl/</td>
</tr>
<tr>
<td>1</td>
<td>song</td>
<td>kesong</td>
<td>/kəˈsɔŋ/</td>
</tr>
<tr>
<td>1</td>
<td>sorry</td>
<td>dedsorry</td>
<td>/dɛdˈsɔrɪ/</td>
</tr>
<tr>
<td>1</td>
<td>special</td>
<td>vespecial</td>
<td>/vəˈspeʃəl/</td>
</tr>
<tr>
<td>1</td>
<td>ten</td>
<td>neten</td>
<td>/nɛˈtɛn/</td>
</tr>
<tr>
<td>1</td>
<td>thousand</td>
<td>nelthousand</td>
<td>/nɛlˈθɑzənd/</td>
</tr>
<tr>
<td>1</td>
<td>wish</td>
<td>ingwish</td>
<td>/ɪŋˈwɪʃ/</td>
</tr>
<tr>
<td>1</td>
<td>wrong</td>
<td>bicwrong</td>
<td>/bɪkˈrɔŋ/</td>
</tr>
<tr>
<td>1</td>
<td>zoo</td>
<td>venzoo</td>
<td>/vənˈzu/</td>
</tr>
<tr>
<td>2</td>
<td>ache</td>
<td>shenache</td>
<td>/ʃənˈeɪk/</td>
</tr>
<tr>
<td>2</td>
<td>axe</td>
<td>dingaxe</td>
<td>/dɪŋˈæks/</td>
</tr>
<tr>
<td>2</td>
<td>breath</td>
<td>shenbreath</td>
<td>/ʃənˈbreθ/</td>
</tr>
<tr>
<td>2</td>
<td>cope</td>
<td>rencope</td>
<td>/rənˈkʊp/</td>
</tr>
<tr>
<td>2</td>
<td>dim</td>
<td>zezdim</td>
<td>/zəzˈdɪm/</td>
</tr>
<tr>
<td>2</td>
<td>dip</td>
<td>vensdip</td>
<td>/vənsˈdɪp/</td>
</tr>
<tr>
<td>2</td>
<td>discipline</td>
<td>gediscipline</td>
<td>/gəˈdɪsəplən/</td>
</tr>
<tr>
<td>2</td>
<td>distant</td>
<td>tezdistant</td>
<td>/təzˈdɪstənt/</td>
</tr>
<tr>
<td>2</td>
<td>effort</td>
<td>sensteffort</td>
<td>/sənstˈɛfət/</td>
</tr>
<tr>
<td>2</td>
<td>emphasis</td>
<td>kelemphasis</td>
<td>/kælˈɛmfaʃəs/</td>
</tr>
<tr>
<td>2</td>
<td>equal</td>
<td>shendequal</td>
<td>/ʃəndˈɪkwəl/</td>
</tr>
<tr>
<td>word</td>
<td>pronunciation</td>
<td>word</td>
<td>pronunciation</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>------</td>
<td>---------------</td>
</tr>
<tr>
<td>fee</td>
<td>tre'fi</td>
<td>fool</td>
<td>jə'ful</td>
</tr>
<tr>
<td>hook</td>
<td>gri'hɔk</td>
<td>hop</td>
<td>gres'hɔp</td>
</tr>
<tr>
<td>leather</td>
<td>dəs'leðə</td>
<td>mirror</td>
<td>tra'mɪɾə</td>
</tr>
<tr>
<td>orient</td>
<td>prɛ'tɔriənt</td>
<td>salary</td>
<td>pə'sæləri</td>
</tr>
<tr>
<td>shot</td>
<td>pən'ʃot</td>
<td>shower</td>
<td>ɡə'ʃoʊə</td>
</tr>
<tr>
<td>stem</td>
<td>dəd'stɛm</td>
<td>strength</td>
<td>lɪŋ'strɛŋθ</td>
</tr>
<tr>
<td>threat</td>
<td>dən'θrɛt</td>
<td>tremble</td>
<td>spə'trembəl</td>
</tr>
</tbody>
</table>
## PART II – Frequency level 3

### Initial Position

#### Difficult Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>ace</td>
<td>acewuck</td>
<td>əɪswʌk</td>
</tr>
<tr>
<td>3</td>
<td>apt</td>
<td>aptob</td>
<td>æptob</td>
</tr>
<tr>
<td>3</td>
<td>bang</td>
<td>bangthot</td>
<td>ˈbæŋθɔt</td>
</tr>
<tr>
<td>3</td>
<td>canoe</td>
<td>canoefrab</td>
<td>kəˈnufræb</td>
</tr>
<tr>
<td>3</td>
<td>complexion</td>
<td>complexiontane</td>
<td>kəmˈplɛkʃənteiŋ</td>
</tr>
<tr>
<td>3</td>
<td>conception</td>
<td>conceptionigs</td>
<td>kənˈsɛpʃənɪɡz</td>
</tr>
<tr>
<td>3</td>
<td>confession</td>
<td>confessionid</td>
<td>kənˈfɛʃənid</td>
</tr>
<tr>
<td>3</td>
<td>destine</td>
<td>destinezeng</td>
<td>ˈdestænzən</td>
</tr>
<tr>
<td>3</td>
<td>elegant</td>
<td>elegantesp</td>
<td>ˈɛləɡəntɛsp</td>
</tr>
<tr>
<td>3</td>
<td>epidemic</td>
<td>epidemicwug</td>
<td>əˈpɛdɪmɪkwʌɡ</td>
</tr>
<tr>
<td>3</td>
<td>fret</td>
<td>fretez</td>
<td>ɛfretɛz</td>
</tr>
<tr>
<td>3</td>
<td>happily</td>
<td>happilyblaf</td>
<td>ˈhæpəliblæf</td>
</tr>
<tr>
<td>3</td>
<td>hearth</td>
<td>hearthruv</td>
<td>hæθrʌv</td>
</tr>
<tr>
<td>3</td>
<td>horror</td>
<td>horrorflen</td>
<td>ˈhɔrɔflɛn</td>
</tr>
<tr>
<td>3</td>
<td>lounge</td>
<td>loungeull</td>
<td>lʌŋdʒʊl</td>
</tr>
<tr>
<td>3</td>
<td>monopoly</td>
<td>monopolybrab</td>
<td>məˈnɒpəlɪbræb</td>
</tr>
<tr>
<td>3</td>
<td>oat</td>
<td>oatrop</td>
<td>əʊtrɒp</td>
</tr>
<tr>
<td>3</td>
<td>peasant</td>
<td>peasantra</td>
<td>pɛzəntræ</td>
</tr>
<tr>
<td>3</td>
<td>penguin</td>
<td>penguinoom</td>
<td>pɛŋgwənʊm</td>
</tr>
<tr>
<td>3</td>
<td>pessimist</td>
<td>pessimistalv</td>
<td>pɛzəmɪstəlɛv</td>
</tr>
<tr>
<td>3</td>
<td>pose</td>
<td>posert</td>
<td>pʊozɔt</td>
</tr>
<tr>
<td>3</td>
<td>rot</td>
<td>rotum</td>
<td>rɒtʌm</td>
</tr>
<tr>
<td>3</td>
<td>superb</td>
<td>superbix</td>
<td>səˈpɜbɪks</td>
</tr>
<tr>
<td>3</td>
<td>suspicious</td>
<td>suspiciousig</td>
<td>səˈspɪʃəsɪɡ</td>
</tr>
<tr>
<td>3</td>
<td>tag</td>
<td>taglems</td>
<td>tæglɛmz</td>
</tr>
<tr>
<td>3</td>
<td>wag</td>
<td>wagree</td>
<td>wægri</td>
</tr>
<tr>
<td>3</td>
<td>wrath</td>
<td>wrathip</td>
<td>rɒθɪp</td>
</tr>
</tbody>
</table>
## Easy Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>abolish</td>
<td>abolishpow</td>
<td>əˈbolɪpɔʊ</td>
</tr>
<tr>
<td>3</td>
<td>affection</td>
<td>affectionren</td>
<td>əˈfɛkʃən rin</td>
</tr>
<tr>
<td>3</td>
<td>ballot</td>
<td>ballotnim</td>
<td>bælətnim</td>
</tr>
<tr>
<td>3</td>
<td>cherish</td>
<td>cherishprem</td>
<td>tʃərɪprem</td>
</tr>
<tr>
<td>3</td>
<td>cling</td>
<td>clingeems</td>
<td>klinjimz</td>
</tr>
<tr>
<td>3</td>
<td>condemn</td>
<td>condemngy</td>
<td>kənˈdemɡi</td>
</tr>
<tr>
<td>3</td>
<td>cosmopolitan</td>
<td>cosmopolitanfict</td>
<td>kəʊzməˈplətnfikt</td>
</tr>
<tr>
<td>3</td>
<td>crept</td>
<td>creptmots</td>
<td>krepmtɔts</td>
</tr>
<tr>
<td>3</td>
<td>earl</td>
<td>earlpud</td>
<td>ɛəlpud</td>
</tr>
<tr>
<td>3</td>
<td>emblem</td>
<td>emblemrels</td>
<td>ɛmblɛrmɛls</td>
</tr>
<tr>
<td>3</td>
<td>exert</td>
<td>exertnol</td>
<td>əɡˈzætnol</td>
</tr>
<tr>
<td>3</td>
<td>hood</td>
<td>hoodgugs</td>
<td>həʊdɡʌz</td>
</tr>
<tr>
<td>3</td>
<td>ignorant</td>
<td>ignorantlem</td>
<td>ɪɡnərəntlɛm</td>
</tr>
<tr>
<td>3</td>
<td>mimic</td>
<td>mimicvops</td>
<td>mɪmɪkvɔps</td>
</tr>
<tr>
<td>3</td>
<td>nymph</td>
<td>nymphwab</td>
<td>nɪmfwæb</td>
</tr>
<tr>
<td>3</td>
<td>obstinate</td>
<td>obstinateklosh</td>
<td>əʊbstenətklɒʃ</td>
</tr>
<tr>
<td>3</td>
<td>oriental</td>
<td>orientalpab</td>
<td>əˈrɪəntlæb</td>
</tr>
<tr>
<td>3</td>
<td>pope</td>
<td>popeteft</td>
<td>pəʊpetɛft</td>
</tr>
<tr>
<td>3</td>
<td>premise</td>
<td>premisethal</td>
<td>prɛmɪs@θæl</td>
</tr>
<tr>
<td>3</td>
<td>rag</td>
<td>ragdist</td>
<td>ræɡdist</td>
</tr>
<tr>
<td>3</td>
<td>skim</td>
<td>skimlurg</td>
<td>skɪmlɜɡ</td>
</tr>
<tr>
<td>3</td>
<td>sob</td>
<td>sobfay</td>
<td>soʊfei</td>
</tr>
<tr>
<td>3</td>
<td>specimen</td>
<td>specimenlang</td>
<td>spɛsəmənˈlæŋ</td>
</tr>
<tr>
<td>3</td>
<td>sped</td>
<td>spedgix</td>
<td>spɛdɡɪks</td>
</tr>
<tr>
<td>3</td>
<td>wed</td>
<td>wednop</td>
<td>wɛdnoʊp</td>
</tr>
</tbody>
</table>
## Stress Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>applicant</td>
<td>applicanten</td>
<td>æpləkəntən</td>
</tr>
<tr>
<td>3</td>
<td>babble</td>
<td>babbleige</td>
<td>bæbləidʒ</td>
</tr>
<tr>
<td>3</td>
<td>cosmic</td>
<td>cosmicent</td>
<td>kɔzmɪkənt</td>
</tr>
<tr>
<td>3</td>
<td>crisp</td>
<td>crispted</td>
<td>krisptəd</td>
</tr>
<tr>
<td>3</td>
<td>desolate</td>
<td>desolatement</td>
<td>dɛsəlatəmɛnt</td>
</tr>
<tr>
<td>3</td>
<td>ebb</td>
<td>ebbgend</td>
<td>ɛbdʒənd</td>
</tr>
<tr>
<td>3</td>
<td>empress</td>
<td>empressesh</td>
<td>ɛmprəsəʃ</td>
</tr>
<tr>
<td>3</td>
<td>folly</td>
<td>follyzek</td>
<td>fɔlɪzək</td>
</tr>
<tr>
<td>3</td>
<td>harsh</td>
<td>harshweld</td>
<td>hæʃwɛld</td>
</tr>
<tr>
<td>3</td>
<td>immigrant</td>
<td>immigrantble</td>
<td>ɪmɪɡrəntbl</td>
</tr>
<tr>
<td>3</td>
<td>leapt</td>
<td>leaptem</td>
<td>lɛptəm</td>
</tr>
<tr>
<td>3</td>
<td>linger</td>
<td>lingerbalp</td>
<td>lɪŋɡəbalp</td>
</tr>
<tr>
<td>3</td>
<td>orphan</td>
<td>orphanlen</td>
<td>ɔfənlen</td>
</tr>
<tr>
<td>3</td>
<td>perish</td>
<td>perishent</td>
<td>ˈpərɪʃənt</td>
</tr>
<tr>
<td>3</td>
<td>smog</td>
<td>smogbent</td>
<td>ɔməɡbənt</td>
</tr>
<tr>
<td>3</td>
<td>swamp</td>
<td>swampent</td>
<td>ɔswɔmpɛnt</td>
</tr>
<tr>
<td>3</td>
<td>tardy</td>
<td>tardymed</td>
<td>tədɪmɛd</td>
</tr>
<tr>
<td>3</td>
<td>torrent</td>
<td>torrentej</td>
<td>tɔrɛntɛdʒ</td>
</tr>
<tr>
<td>3</td>
<td>yacht</td>
<td>yachtsesh</td>
<td>ʃɔtɛʃ</td>
</tr>
</tbody>
</table>
## Final Position

### Difficult Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>algebra</td>
<td>fisalgebra</td>
<td>/ˌfɪsəlˈdʒæbrə/</td>
</tr>
<tr>
<td>3</td>
<td>caress</td>
<td>kēscaˈrēs</td>
<td>/ˈkēskəˈrēs/</td>
</tr>
<tr>
<td>3</td>
<td>comic</td>
<td>voscomic</td>
<td>/ˈvɒskəmɪk/</td>
</tr>
<tr>
<td>3</td>
<td>competence</td>
<td>nescompetence</td>
<td>/ˈneskəmpətəns/</td>
</tr>
<tr>
<td>3</td>
<td>confess</td>
<td>glisconfess</td>
<td>/ˈɡlɪskənfɛs/</td>
</tr>
<tr>
<td>3</td>
<td>eloquent</td>
<td>temelquent</td>
<td>/ˈteməlkwənt/</td>
</tr>
<tr>
<td>3</td>
<td>ignorance</td>
<td>besignorance</td>
<td>/ˌbɛsɪgnərəns/</td>
</tr>
<tr>
<td>3</td>
<td>lease</td>
<td>bablease</td>
<td>/ˈbæblɪs/</td>
</tr>
<tr>
<td>3</td>
<td>lick</td>
<td>Beplick</td>
<td>/ˈbɛplɪk/</td>
</tr>
<tr>
<td>3</td>
<td>malice</td>
<td>deesmalice</td>
<td>/dɪzmæləs/</td>
</tr>
<tr>
<td>3</td>
<td>mob</td>
<td>fɛsmob</td>
<td>/fɛsmoʊb/</td>
</tr>
<tr>
<td>3</td>
<td>mole</td>
<td>kɛsmole</td>
<td>/kɛsmɔʊl/</td>
</tr>
<tr>
<td>3</td>
<td>nimble</td>
<td>wɪsnimbɛl</td>
<td>/wɪznimbɛl/</td>
</tr>
<tr>
<td>3</td>
<td>obstacle</td>
<td>fɔʃobstacle</td>
<td>/fəˈɔjobstækəl/</td>
</tr>
<tr>
<td>3</td>
<td>parish</td>
<td>nɛsparish</td>
<td>/nɛspærɪʃ/</td>
</tr>
<tr>
<td>3</td>
<td>peril</td>
<td>gursperil</td>
<td>/ˈɡɜrsperɪl/</td>
</tr>
<tr>
<td>3</td>
<td>perplex</td>
<td>θɔspəˈplɛks</td>
<td>/θɔspəˈplɛks/</td>
</tr>
<tr>
<td>3</td>
<td>pheasant</td>
<td>gospeasant</td>
<td>/ɡəʊspəˈzɛnt/</td>
</tr>
<tr>
<td>3</td>
<td>rack</td>
<td>gedræk</td>
<td>/ɡɛdˈræk/</td>
</tr>
<tr>
<td>3</td>
<td>relish</td>
<td>fligrelish</td>
<td>/ˈflɪɡrɛlɪʃ/</td>
</tr>
<tr>
<td>3</td>
<td>remembrance</td>
<td>mekremembrance</td>
<td>/mɛkˈrɛmənbrəns/</td>
</tr>
<tr>
<td>3</td>
<td>resemblance</td>
<td>sɨθrəˈzembləns</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>rip</td>
<td>sæbrip</td>
<td>/sæbrip/</td>
</tr>
<tr>
<td>3</td>
<td>wrestle</td>
<td>vibwrestle</td>
<td>/vɪbˈrɛstəl/</td>
</tr>
</tbody>
</table>
## Easy Condition

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Target Word</th>
<th>Training Word</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>bishop</td>
<td>plutbishop</td>
<td>/plʌtbɪʃ/</td>
</tr>
<tr>
<td>3</td>
<td>blouse</td>
<td>ostblouse</td>
<td>/ɒst'blauz/</td>
</tr>
<tr>
<td>3</td>
<td>bull</td>
<td>pedbull</td>
<td>/pɛd'bʊl/</td>
</tr>
<tr>
<td>3</td>
<td>compel</td>
<td>kedcompel</td>
<td>/ke'dkəm'pɛl/</td>
</tr>
<tr>
<td>3</td>
<td>conjecture</td>
<td>gipconjecture</td>
<td>/ɡɪp'kɒn'dɛktʃə/</td>
</tr>
<tr>
<td>3</td>
<td>cosmos</td>
<td>fropcosmos</td>
<td>/frɒ'kɒzmos/</td>
</tr>
<tr>
<td>3</td>
<td>decree</td>
<td>sugdecree</td>
<td>/sæg'de'kri/</td>
</tr>
<tr>
<td>3</td>
<td>desert (verb)</td>
<td>yagdesert</td>
<td>/ɔæɡ'de'zæt/</td>
</tr>
<tr>
<td>3</td>
<td>ferry</td>
<td>gickferry</td>
<td>/gɪk'fɛri/</td>
</tr>
<tr>
<td>3</td>
<td>geese</td>
<td>rospgeese</td>
<td>/rɔsp'ɡiːz/</td>
</tr>
<tr>
<td>3</td>
<td>lurk</td>
<td>fimlurk</td>
<td>/fɪmlək/</td>
</tr>
<tr>
<td>3</td>
<td>magician</td>
<td>noltmagician</td>
<td>/nɔltma'dʒɪʃən/</td>
</tr>
<tr>
<td>3</td>
<td>mock</td>
<td>goltmock</td>
<td>/ɡɔltmək/</td>
</tr>
<tr>
<td>3</td>
<td>mustache</td>
<td>sefmustache</td>
<td>/sefməstɑːʃ/</td>
</tr>
<tr>
<td>3</td>
<td>nominal</td>
<td>leknominal</td>
<td>/leknə'mənəl/</td>
</tr>
<tr>
<td>3</td>
<td>pang</td>
<td>voadpang</td>
<td>/vɔodpæŋ/</td>
</tr>
<tr>
<td>3</td>
<td>peck</td>
<td>stidpeck</td>
<td>/stɪd'pɛk/</td>
</tr>
<tr>
<td>3</td>
<td>perception</td>
<td>agperception</td>
<td>/æɡpə'sɛpʃən/</td>
</tr>
<tr>
<td>3</td>
<td>pest</td>
<td>doikpest</td>
<td>/dɔɪk'pɛst/</td>
</tr>
<tr>
<td>3</td>
<td>physician</td>
<td>preetphysician</td>
<td>/prɪt'fɛzɪʃən/</td>
</tr>
<tr>
<td>3</td>
<td>plea</td>
<td>congplea</td>
<td>/kɔŋ'pli/</td>
</tr>
<tr>
<td>3</td>
<td>prohibit</td>
<td>pidprohibit</td>
<td>/pɪdprə'hɪbət/</td>
</tr>
<tr>
<td>3</td>
<td>reap</td>
<td>huvreap</td>
<td>/hʌvri)p</td>
</tr>
<tr>
<td>3</td>
<td>remnant</td>
<td>foonremnant</td>
<td>/fɔn'remənt/</td>
</tr>
<tr>
<td>3</td>
<td>resemble</td>
<td>kelresemble</td>
<td>/kɛlre'zəmbəl/</td>
</tr>
<tr>
<td>3</td>
<td>theft</td>
<td>pebtheft</td>
<td>/pɛbθɛft/</td>
</tr>
<tr>
<td>3</td>
<td>trolley</td>
<td>waptrolley</td>
<td>/wæptrəli/</td>
</tr>
<tr>
<td>3</td>
<td>turf</td>
<td>gedturf</td>
<td>/ɡɛd'tɔːf/</td>
</tr>
<tr>
<td>3</td>
<td>zest</td>
<td>bep zest</td>
<td>/bɛp'zɛst/</td>
</tr>
<tr>
<td>Frequency</td>
<td>Target Word</td>
<td>Training Word</td>
<td>IPA</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>---------------</td>
<td>-----------</td>
</tr>
<tr>
<td>3</td>
<td>alley</td>
<td>spenalley</td>
<td>spən'æli</td>
</tr>
<tr>
<td>3</td>
<td>auction</td>
<td>beledauction</td>
<td>bəl'dәkʃәn</td>
</tr>
<tr>
<td>3</td>
<td>baffle</td>
<td>gebaffle</td>
<td>ɡə'bæfel</td>
</tr>
<tr>
<td>3</td>
<td>chap</td>
<td>nelchap</td>
<td>nəl'tʃæp</td>
</tr>
<tr>
<td>3</td>
<td>competent</td>
<td>beocompetent</td>
<td>bә'kөmpәtәnt</td>
</tr>
<tr>
<td>3</td>
<td>eloquence</td>
<td>teleloquence</td>
<td>tәl'әlәkwәns</td>
</tr>
<tr>
<td>3</td>
<td>ferry</td>
<td>shenferry</td>
<td>fәn'fәri</td>
</tr>
<tr>
<td>3</td>
<td>germ</td>
<td>telgerm</td>
<td>tәl'dʒәm</td>
</tr>
<tr>
<td>3</td>
<td>gown</td>
<td>gezgown</td>
<td>gәz'ɡәn</td>
</tr>
<tr>
<td>3</td>
<td>heather</td>
<td>ningheather</td>
<td>nɪŋ'hәðә</td>
</tr>
<tr>
<td>3</td>
<td>hoof</td>
<td>belhoof</td>
<td>bәl'hof</td>
</tr>
<tr>
<td>3</td>
<td>jerk</td>
<td>trejerk</td>
<td>trә'dʒәk</td>
</tr>
<tr>
<td>3</td>
<td>lag</td>
<td>linglag</td>
<td>lɪŋ'lәɡ</td>
</tr>
<tr>
<td>3</td>
<td>moth</td>
<td>desmoth</td>
<td>dәs'mәθ</td>
</tr>
<tr>
<td>3</td>
<td>oath</td>
<td>mingoath</td>
<td>mɪŋ'oθ</td>
</tr>
<tr>
<td>3</td>
<td>ominous</td>
<td>jelominous</td>
<td>dʒәl'әmәnәs</td>
</tr>
<tr>
<td>3</td>
<td>orator</td>
<td>tingorator</td>
<td>tɪŋ'әrәtә</td>
</tr>
<tr>
<td>3</td>
<td>oyster</td>
<td>tendoyster</td>
<td>tәn'dәʊәstә</td>
</tr>
<tr>
<td>3</td>
<td>post</td>
<td>kepост</td>
<td>kә'poʊәst</td>
</tr>
<tr>
<td>3</td>
<td>preface</td>
<td>spentpreface</td>
<td>spәnt'prefәs</td>
</tr>
<tr>
<td>3</td>
<td>preference</td>
<td>bepreference</td>
<td>bә'prәfәrәns</td>
</tr>
<tr>
<td>3</td>
<td>prosper</td>
<td>klenproser</td>
<td>kлаn'prospa</td>
</tr>
<tr>
<td>3</td>
<td>sip</td>
<td>bensip</td>
<td>ɓәn'sip</td>
</tr>
<tr>
<td>3</td>
<td>skeleton</td>
<td>senskeleton</td>
<td>sәn'skeләtn</td>
</tr>
<tr>
<td>3</td>
<td>wag</td>
<td>gewag</td>
<td>ɡә'wәәɡ</td>
</tr>
</tbody>
</table>
The IELTS 9-Band Scale

(As given on the official IELTS website (IELTS band scores, n.d.)).

Each band corresponds to a level of English competence. All parts of the test and the Overall Band Score can be reported in whole and half bands, eg 6.5, 7.0, 7.5, 8.0.

**Band 9: Expert user:** has fully operational command of the language: appropriate, accurate and fluent with complete understanding.

**Band 8: Very good user:** has fully operational command of the language with only occasional unsystematic inaccuracies and inappropriacies. Misunderstandings may occur in unfamiliar situations. Handles complex detailed argumentation well.

**Band 7: Good user:** has operational command of the language, though with occasional inaccuracies, inappropriacies and misunderstandings in some situations. Generally handles complex language well and understands detailed reasoning.

**Band 6: Competent user:** has generally effective command of the language despite some inaccuracies, inappropriacies and misunderstandings. Can use and understand fairly complex language, particularly in familiar situations.

**Band 5: Modest user:** has partial command of the language, coping with overall meaning in most situations, though is likely to make many mistakes. Should be able to handle basic communication in own field.

**Band 4: Limited user:** basic competence is limited to familiar situations. Has frequent problems in understanding and expression. Is not able to use complex language.
**Band 3: Extremely limited user:** conveys and understands only general meaning in very familiar situations. Frequent breakdowns in communication occur.

**Band 2: Intermittent user:** no real communication is possible except for the most basic information using isolated words or short formulae in familiar situations and to meet immediate needs. Has great difficulty understanding spoken and written English.

**Band 1: Non-user:** essentially has no ability to use the language beyond possibly a few isolated words.

**Band 0: Did not attempt the test:** No assessable information provided.
Participant Questionnaire

Participant Code: ______

Circle all the words that are true for you.

“The task was...”

too easy  easy  difficult  too difficult

fun  interesting  boring

helpful  confusing

too long  long  short  too short

tiring  exciting
Dictation Texts

Transcript One

Dangers of the Beach (adapted from Salter, 2003).

The typical attitude of Australians to the beach is almost love. They adore the beach. We are a beach culture in Australia. The reason for this is that most Australians live close to the coast (all the Australian cities are by the sea) and Australia is a hot country.

Why do Australians sunbathe? The answer to that is simple; fashion. They think they look good if they’re brown. The result of this has been that Australia has the highest rate of skin cancer in the world. When I was a kid I used to get so badly sunburnt at the beginning of summer that I could hardly move, and then after a few days my skin would start peeling off like a lizard. The government knows about this problem, and the Cancer Council has a campaign to educate Australians about the dangers of sun and skin cancer. The Cancer Council used to have a slogan – it was ‘slip, slop and slap’. Each of these words has a special meaning. All three words mean ‘put on’, but putting on something different. Slip means slip on a shirt, slop means slop on some sunscreen, and slap? Slap on a hat. These three things should protect you from the sun.
Transcript Two

Shark Feeder (Brawn, 2002)

Teacher: Um okay. Attention, everybody. Thank you. This is Sam Jacobson. He's a shark feeder at the Aquarium.

Sam: Hello kids.

Student: Hello Sam. [GIGGLING]

Teacher: Now Sam has been, um, very generous with his time today. He's going to talk to you all about his job. Thank you, Sam, for giving up your time today. Um, could you tell us all what do shark feeders do?

Sam: Well, we feed the fish primarily and we also hop in and clean the tank.

Student: With the sharks?

Sam: [LAUGHS] Yeah, we, ah, we wear chain mail gloves to protect our hands.

Teacher: And what exactly is it that you do each day?

Sam: Well, basically we start the day with food preparation. First we defrost the frozen fish and the squids from the market. Then we get the feed tubs ready. Next we pop our wetsuits on and then we start the feeding. The big stingrays are first and so on and so on and, ah, we finish up lastly with the sharks. We hand feed the sharks. We don't feed them with a stick. We feed the sharks from the front because they can turn their heads very quickly to the side and ah, they could probably, you know, bite your hand, or they could rip it off.
Transcript Three

A Car Accident (Salter, 2003)

Leigh: In 1974 I went travelling with a friend of mine.

Alicia: Where did you go?

Leigh: We went to a place called Wilson’s Promontory – that’s about 200 kilometres from Melbourne. We were travelling in an old VW Beetle, 1962 model. This area is well known because it’s the most southerly point on the Australian mainland. It’s got great scenery and beaches as well, and not many big trees, but lots of small bushes – very thick they are. We were driving along and we saw a sign which said there was a beach, so we wanted to check that out. We turned off the road. We were in a lot of danger on this road.

Alicia: Why was that?

Leigh: Well the road was narrow, the sun was very low (it was shining in our eyes), but as usual, the worst problem – we were driving much too fast. Suddenly the road finished in a car park. I turned to Pete and said, “Where’s the road?” Then I put the brakes on and skidded right across the car park, deep into the scrub – about ten metres in.

Alicia: So what happened to the car?

Leigh: The car had a bit of damage: it had a dent on the front guard, one headlight was broken and there were lots of scratches down both sides.
Transcript Four

Obesity in Children (Brawn, 2002)

Kerry: To today’s program: Overweight Kids. According to medical research, obesity is fast becoming the most serious and costly disease among our children. Overeating and under-exercising lie at the heart of the problem. Now to discuss what lies behind the obesity epidemic and what treatment can help to stem the tide, I’m joined by Dr Thomas Gold.

Thomas: Good morning Kerry.

Kerry: Welcome to the show, Thomas. The current figures on obesity among children and the adult population are quite alarming. What are the reasons for the increase?

Thomas: Well, Kerry. They’re, they’re certainly alarming. I think we’re all aware of the changes in society which have reduced our physical activity and changed our eating habits. Ah, we don’t do much physical labour at work anymore. In fact most of us sit down in front of our computers for eight hours a day.

Kerry: [LAUGHS] That’s true.

Thomas: We do absolutely nothing. Most of us drive and very few of us walk, we don’t allow, um, our children to cycle or to play after school.

Kerry: Oh, well, there’s the safety issue.

Thomas: We’re very concerned about safety. We drag them inside. We sit them down. We don’t allow them to exercise.
Transcript Five

Hospitality Class (Brawn, 2002)

Okay. Today we’re going to do some more vegetarian dishes. I’m going to talk about some ingredients. What we’ve got here is two types of tofu. This one’s more a dessert tofu. Now we’re going to use the firm tofu. This is a more savoury one. When you touch it, it’s very firm on the outside and that’s made from soybean curd. Okay? It’s quite good for you. It doesn’t have much flavour but it’s good for vegetarians ’cause they want to have a stir fry with some extra... obviously protein things. Okay? ’Cause there’s obviously a lot in soy bean. Hopefully it hasn’t been genetically altered.

Okay? I’d like now to move on the other vegetables. So what have we got? We’ve got baby bok choi over there. Give them a good wash because those little bulbs, you’ll find little grubs with it, so. Also we’ve got Shitaki mushrooms. These are still imported from China. Don’t ask me why, but we import them from China. I don’t know why we can’t grow them here. You’ve got to take the stalk out because it’s very tough, but it’s still nicer than the dry ones and it’s very, very strong in flavour.
Transcript Six

Indian Juice Story (Salter, 2003)

Jim: Leigh, tell me one of your Indian stories.

Leigh: Sure Jim. In 1986 I went to the north-west of India, to the state of Rajasthan. There’s a small town there called Pushka, which is very well known. It’s a desert town and it’s best known outside India for its camel fair – that’s when hundreds of camels are brought into town and there’s a huge camel market. This is in the desert, and the days are very hot and the nights are cool. It’s very low humidity there.

I went to Pushka for a few reasons. I’d read about it, and I knew that it was a religious centre (Hindus from all over India come there to pray). It’s also a very old town and there are some interesting old buildings - and the other thing I liked was the monkeys. Monkeys live wild around here because, according to the Hindus, the monkey is a holy animal – and they’re interesting to watch.

The monkeys were a big problem in this town. They used to steal. They were very cheeky, and they’d come down from the trees to the hotels where the tourists were staying, take clothes, cameras and other things, and run back to the treetops with them.
Transcript Seven

Crime Buster (Brawn, 2002)

Tony: Good morning. It’s Tony Chan on Facets. Today on our program our
guest is Senior Constable Olympia Demos to talk about career and
training for the police force. Good morning, Olympia. Welcome to
the program.

Olympia: Thank you, Tony.

Tony: What sort of services are the police offering these days?

Olympia: Well, the police offer so many services now; it’s almost hard to list
them - from investigating major crimes such as murders, investigating
accidents, to taking reports of stealing, to getting the cat from up the
tree, so to speak.

Tony: Mm hm.

Olympia: We have a lot of neighbourhood disputes we have to deal with. I, I
mean, the list goes on. There’s domestic-related issues that police
have to intervene in to protect people. The law enforcement aspect of
policing probably takes quite a small part of the day compared to the
other customer service issues. You just don’t know when you start the
day as a duty officer what you’re gonna end up with.

Tony: So what do you need to be able to do as a police officer?

Olympia: Well, um, peacekeeper, yeah? Trying to negotiate between parties that
have differences. You have to be a fairly confident person to be in the
service, fairly mature and fairly assertive. And you also need to be able
to cope with, um, well you need a fair amount of patience because
there’s a bit of paperwork involved.
Transcript Eight

A Bicycle Accident (adapted from Salter, 2003)

I’d like to tell you about a terrible accident which happened to me once, in February ’92 in Perth. I had a very serious accident on the cycle path beside the Swan River, opposite the Burswood Casino. The weather conditions were typical for that time of year: summer, cloudless blue sky, and the Fremantle Doctor was blowing.

The Fremantle Doctor is a wind. It’s the name of a wind which comes in summer when the weather is very hot, and this wind comes from the direction of Freemantle, south of Perth, and it’s called the Doctor because it fixes the heat – after the wind comes in it gets very cool.

So anyway, I was zooming along the cycle path after school one day. I was travelling as fast as I could, and with the wind behind me, a tail wind, I think I was going about 40 kilometres an hour. I was pedalling as fast as possible. I heard a woman shouting a name, like she was calling a child, and then a few seconds later, a Border collie (black and white dog) ran up right beside me on my bicycle. I looked at it. I didn’t have time to think or do anything, when the dog ran straight in front of my bike.
Transcript Nine

How to Make Gazpacho Soup (Brawn, 2002)

So today we’re going to make a few small dishes. We’re going to make this one first. It’s gazpacho soup. What we do, um, I’ve pretty well written the method up there on the board. So we’ve got everything here. To begin with, we’ve got the cut cucumber, tomatoes, which have been peeled and de-seeded, onion, garlic, stock, olive oil, salt and pepper. And the only thing I haven’t got in there is vinegar.

First I’m going to blend the tomatoes, cucumber, onion and garlic up into a mixture. It looks this pink colour because we’re using fresh tomato. You might have seen the recipe in other books. They use tinned tomato, that’s why it’s a darker, redder colour. Next add stock and part of the olive oil, not too much, about twelve mls, OK. And then it’s going to have the garnish of the capsicum and the other half of the cucumber. You taste that and see if you need any more salt and pepper.

OK and turning now to the croutons. Now we’re going to make the croutons, OK. So with the croutons, first you just cut the crust off the bread. Cut them into six by six long way, six across - and dry fry them in the oven, no butter, no oil, no nothing, OK? Next put the croutons over the top. Add some fresh basil there. I’m going to put a little bit over the top.
Participant Questionnaire

We compared three types of training that were intended to help you listen to speech.

1. Which type was the MOST helpful?

Circle one: dictation  connected speech training  word-spotting

Why?
____________________________________________________________________

____________________________________________________________________

2. Which type was the LEAST helpful?

Circle one: dictation  connected speech training  word-spotting

Why?
____________________________________________________________________

____________________________________________________________________

How do you think this training could be improved?
____________________________________________________________________

____________________________________________________________________
This questionnaire refers to the **dictation training**.

Circle all the words that are true for you.

“The training was...”

- too easy
- easy
- difficult
- too difficult
- fun
- interesting
- boring
- helpful
- confusing
- too long
- long
- short
- too short
- tiring
- exciting

Or add your own ideas: ________________________________________________

____________________________________________________________________

____________________________________________________________________

____________________________________________________________________
This questionnaire refers to the connected speech training.

Circle all the words that are true for you.

“The training was...”

too easy          easy           difficult          too difficult

fun              interesting       boring

helpful          confusing

too long          long            short            too short

tiring            exciting

Or add your own ideas: ____________________________________________
___________________________________________________________________
___________________________________________________________________
___________________________________________________________________
This questionnaire refers to the **word-spotting training**.

Circle all the words that are true for you.

“The training was…”

- too easy
- easy
- difficult
- too difficult

- fun
- interesting
- boring

- helpful
- confusing

- too long
- long
- short
- too short

- tiring
- exciting

Or add your own ideas: _______________________________________________
__________________________________________________________
_________________________________________________________________
_________________________________________________________________
Experiment One Results

See Chapter 5 for further details.

5.4.2 Pre-test - Post-Test 1

There was a significant main effect of ‘testing time’ from the pre-test to post-test 1, $F(1, 20) = 5.65, p = .03$. There was no significant main effect of ‘training group’, $F(2, 20) = 2.10, p = .15$. The interaction between ‘training group’ and ‘testing time’ was not significant, $F(2, 20) = 1.22, p = .32$.

Results Groups

An analysis of the three different results groups (High, Mid and Low according to their pre-test scores) across the pre-test and post-test 1 showed that the main effect of ‘testing time’ from pre-test to post-test 1 was not significant, $F(1, 20) = 2.40, p = .14$. However, the ‘results group’ by ‘testing time’ interaction was significant, $F(2, 20) = 3.44, p = .05$, as was the main effect of ‘results group’, $F(2, 20) = 19.89, p < .001$, with the mean scores of the groups significantly different in the direction High>Mid>Low at the $p = <.001$ level, except for Mid>Low, $p = .007$.

There was not a significant interaction for ‘training group’ and ‘testing time’ within the ‘results group’ analyses, $F(2, 20) = 1.19, p = .35$.

5.4.3 Pre-Test - Post-Test 2

There was a significant effect of ‘testing time’, $F(2, 17) = 6.62, p = .02$, but no ‘testing time’ by ‘training group’ interaction, $F(2, 17) = .35, p = .71$. 
Results Groups

The main effect of 'testing time' from pre-test to post-test 2 for the results groups was not significant, $F(2, 17) = 2.63, p = .17$. The 'results group' by 'testing time' interaction was also not significant, $F(2, 17) = .22, p = .81$. There was no significant interaction for 'training group' and 'testing time' within the 'results group' analyses (Testing time x Results Group x Training Group), $F(2, 17) = 1.72, p = .22$.

5.4.4 Pre-Test - Post-Test 3

The main effect of 'testing time' from pre-test to post-test 3, $F(1, 14) = 2.42, p = .14$ and 'training group' $F(2, 14) = 2.65, p = .11$, were not significant. The 'training group' by 'testing time' interaction was also not significant $F(2, 14) = 1.65, p = .23$.

Results Groups

An analysis based on the 'results groups' (High, Mid, Low depending on their pre-test scores) was not conducted due to unbalanced participant numbers (High, $n = 2$, Mid, $n = 12$ and Low, $n = 2$).

5.4.5 Pre-Test and All Post-Tests

Across the four tests there was a significant effect of 'testing time', $F(3, 13) = 5.47, p = .004$. A simple pairwise comparison shows that there was a significant difference between scores on post-test 1 and post-test 2, with post-test 1 scores 7.78% higher than post-test 2, $p = .04$, and post-test 1 and post-test 3, with mean scores for post-test 1 being 8.94% higher than post-test 3, $p = .02$. However, there was no significant effect of 'training group' $F(2, 13) = 1.69, p = .23$, or of interaction of 'testing time' by 'training group', $F(6, 13) = .79, p = .58$. 

304
Listening Cloze Pre-test Sentence List

The sentences are organised by their origin. The ‘bleeped out’ word is underlined and its part of speech listed.

PART I: Sentences adapted and recorded by the author.

A: Adapted from Akker and Cutler (2003)

1. N: The villa with the carport must belong to the doctor's widow.

2. V: The actions of the crew focused on the dangerous situation.

3. ADJ: The statement of the crown witness led to the burglar's arrest.

4. N: The mother of two daughters wrote to the boarding school.

5. V: The bones of the dinosaur were found by the Cuban archaeologist.

6. ADJ: The rising price of boxes worried the dog food manufacturer.

B: Adapted from Cutler and Darwin (1981)

1. N: She managed to remove the dirt from the rug, but not the grass stains.

2. V: The head cashier said the bag contained money.

3. ADJ: John decided that he would take the boat down the river.

4. N: The top experts were all unable to break the code the spy had used.

5. V: The agent was not surprised to find a bug hidden in his telephone.

6. ADJ: She felt very guilty about the debt she owed him.

7. N: The couple had quarreled over a book they hadn't even read.

8. V: The politician reminded his friend of a bet they had made.

9. ADJ: That summer four years ago I ate roast duck for the first time.

10. N: I had been thinking of buying another coat before you mentioned it.
C: TIMIT Acoustic-Phonetic Continuous Speech Corpus (Garofolo, et al., 1993)

1. N: She had your dark suit in greasy wash water all year.

2. V: Don’t ask me to carry an oily rag like that.

3. ADJ: A boring novel is a superb sleeping pill.

4. ADJ: We saw eight tiny icicles below our roof.

5. N: Each untimely income loss coincided with the breakdown of a heating system part.

6. N: The sermon emphasized the need for affirmative action.

7. V: Kindergarten children decorate their classrooms for all holidays.


1. ADJ: We're glad that the price range is smaller than any of us expected.

2. N: The group asked if I wanted to come along on the barge trip.

3. V: Amongst her friends at university she was considered beautiful.

4. ADJ: John could supply him with the latest draft of his work to proof-read.

5. N: Was it the blue globe that broke when he switched on the light?

6. ADV: The chill wind caused the twins to shiver violently in their threadbare clothes.

7. ADV: The table is badly glued and made so sloppily that it tilts.

8. ADV: He hurled several stones from the bridge whilst I quickly ran for cover.

9. N: Thank goodness it’s Friday and time to go home!

10. ADV: Who says itches are always so tempting to scratch?

11. V: The length of her skirt caused the passers-by to glare.
E. Composed by the author

1. ADJ: It's **hard** to think of answers quickly in English.

2. ADJ: Only **thirteen** more shopping days left until Christmas.

PART II: Pre-recorded sentences

F: Nazzi, Jusczyk, & Johnson (2000)

1. V: This is the first time an international exhibit has **come** to this town.

2. N: The Green **Party** unexpectedly gained strong support from middle class people.

3. ADV: Trade Unions have lost a lot of their influence **during** the past ten years.

4. V: Some more money will be needed **to make** this project succeed.

5. N: The latest **events** have caused an outcry in the international community.

6. ADJ: The committee will meet this afternoon for a **special** debate.

7. ADV: The parents **quietly** crossed the dark room and approached the boy's bed.

8. N: Finding a job is difficult in the present economic climate.

9. V: Nobody noticed when the children **slipped** away just after dinner.

10. ADJ: My grandparents' neighbour is the most **charming** person I know.

11. ADJ: Most **European** banks close extremely early on Friday afternoons.

12. V: The government is **planning** a reform of the educational program.

13. N: The recent rainfall has caused very severe **damage** in the higher valleys.

14. N: A hurricane was announced this **afternoon** on the TV.
Obesity in Children

Kerry: To today’s program: Overweight Kids.

, obesity is fast becoming the most serious and costly disease among our children. Overeating and under-exercising

Now to discuss what lies behind the obesity epidemic and what treatment can help to stem the tide, I’m joined by Dr Thomas Gold.

Thomas: Good morning Kerry.

Kerry: . The current figures on obesity among children and the adult population are quite alarming.

?  

Thomas: Well, Kerry. They’re, they’re certainly alarming.

Ah,

. In fact

Kerry: [LAUGHS] That’s true.

Thomas: We do absolutely nothing. and we don’t allow, um,

.

Kerry: Oh, well, there’s the safety issue.

Thomas: We don’t allow them to, to exercise.
Kerry: Mm hmm...

Thomas: And of course in, in,

Kerry: Yes, in fact we’re seeing diseases grow among children that have traditionally been middle-aged diseases, aren’t we?

Thomas: Mm, we are.

um, such as heart disease, diabetes, high blood pressure.

in children as young as, oh you know, five, five and six.

Mm.

Kerry: And, and

? 

Thomas: er,

, ah , .

Kerry: Mm hm.

Thomas: Ah,
Dictation Sentences

The underlined words were missing on the test paper.

Pre-test Sentences

1. This is the first time an international exhibit has come to this town.
2. The actions of the crew focused on the dangerous situation.
3. She managed to remove the dirt from the rug.
4. She had your dark suit in greasy wash water all year.
5. We’re glad that the price range is smaller than any of us expected.
6. The committee will meet this afternoon for a special debate.
7. The mother of two daughters wrote to the boarding school.
8. The head cashier said the bag contained money.
9. Don’t ask me to carry an oily rag like that.
10. John could supply him with the latest draft of his work.
11. The Green Party unexpectedly gained strong support from middle class people.
12. The rising price of boxes worried the dog food manufacturer.
13. John decided he would take the boat down the river.
14. Call an ambulance for medical assistance.
15. Thank goodness it’s Friday and time to go home.
16. Some more money will be needed to make this project succeed.
17. The remains of the camp were found by the deer hunter.
18. The top experts were all unable to break the code.
20. The length of her skirt caused people to stare.
21. The parents quietly crossed the dark room and approached the boy’s bed.
22. The young man on the corner was wearing the blue hat.
23. She felt very guilty about the debt she owed him.
24. The frightened child was comforted by his big brother.
25. She’ll think of an excuse for why she overslept.
26. Her ability to work without supervision is noteworthy.
27. Not enough attention was paid to the cost before the party.
28. The couple fought over a book they had read.
29. The woman with the bag went into the doctor’s office.
30. Most European banks close extremely early on Friday afternoons.
31. The dirty water ruined the coffee’s flavor.
32. My grandparents’ neighbour is the most charming person I know.
33. I had been thinking about buying another coat before the winter.
34. These jokes have been taken much too far.
35. Military officers are expected to obey government orders.
36. The bath plug is missing, so you’ll have to take a shower.
37. The owner of the bookshop refused to go to the party.
38. The tooth fairy forgot to come when Peter’s tooth fell out.
39. The woman found three letters waiting for her on the piano.
40. Finding a job is difficult in the present economic climate.
Post-test 1 Sentences

1. The city council has decided to renovate the community center.
2. This year’s Chinese celebrations were not nearly as impressive as last years.
3. Anne enjoyed her freedom by going to the zoo and seeing the giant pandas.
4. He grabbed a towel and then answered the phone by the clock.
5. The rebuilding of the city started the very first day after the earthquake.
6. Don’t tell anyone this, but the topic of John’s speech is beginning to annoy me.
7. All of those at the meeting decided to take greater care of the environment.
8. Seven paintings of great value have recently been stolen from the museum.
9. My mother gets angry when I say “yeah” instead of “yes”.
10. The film is aimed at younger audiences who love colourful effects.
11. The young, single woman took good care of three small children.
12. Most of the supporters of the football club had to travel for an entire day.
13. She finally moved to the beach after fifty years of complaining about the cold in the mountains.
14. Unions have lost a lot of their influence during the past ten years.
15. Sporting clubs and organizations can make a lot of money from advertising.
16. Jack tied his shoelace quickly after jumping off the step.
17. The woman over there apparently specialises in plastic surgery.
18. The movie goes on a little bit too long, like so many do these days.
19. Spider webs slowly find their way into every corner.
20. Many new parents struggle with babies who cry constantly and have trouble sleeping.
21. The policewoman from the country joined the peace-keeping forces.
22. On childhood family holidays, we swam freely on the wild coast.

23. The people next door could often be heard singing in the morning.

24. She surprised herself and her family by falling in love with her doctor.

25. Nobody noticed when the children left the house just after dinner.

26. Bill could have sat on the train doing nothing but chose to stand.

27. The library is open every day from 8 AM till 6 PM.

28. The chef lost her sense of smell after being hit by a car.

29. My uncle spent most of his time surrounded by beautiful women.

30. Climate change policies may affect the disadvantaged and other low income groups.

31. There is an attractive market twice a week on the main square of that village.

32. More women over forty are meeting new partners through internet dating.

33. The postman always avoided the house because of the dangerous dog.

34. In traditional houses they expect dinner guests to wear evening clothes.

35. It is common knowledge that the kangaroo carries its baby in its front.

36. My parents’ generation is respected for creating important social change for future generations.

37. For lunch I had about four or five fresh picked mangoes.

38. Ben was a pleasant and modest boy until he went into the army.

39. The rich playboy spent most of his time in his luxury car.

40. The typical attitude of Australians to the beach is almost love.
Post-test 2 Sentences

1. He talked a lot about the past and his failed relationship.
2. I don't think I'll be eating fried chicken with coffee again.
3. Many people want a three bedroom house in the suburbs.
4. The very loud television could be heard from next door.
5. Traditional medicine ruled the medical industry about forty years ago.
6. The gift certificate came in handy on the trip.
7. It seemed to me that love was the only cure for fear.
8. When you make a sale discuss the payment terms.
9. Be prepared for the temperature difference between the mountains and the coast.
10. A young reporter arrived at the door one morning.
11. We must not lose touch with the physical community around us.
12. Lack of access to safe, clean drinking water means disease or death.
13. She remained alone in the park and moved to the steps.
14. Earthquakes have a bigger impact on health than other natural disasters.
15. Voting is the most direct path we have to our politicians.
16. Make sure that your client is comfortable with making Internet transfers.
17. The new oil painting suited the room nicely.
18. People are setting up foundations to give money to good causes.
19. The city's citizens love to shop with many seeing it as recreation.
20. Training for nurses should include caring for the elderly.
21. Only the policeman and I were near enough to hear him.
22. We had a difficult conversation for a while and then she hung up.
23. Both of us were nervous at being suddenly alone together.
24. The centre is both a tourist agency and a retail outlet.
25. The sun was shining and there was a gentle breeze blowing.
26. He put his hands in his pockets and pretended he was bored.
27. We are approaching the end and need to choose our direction.
28. You should work within the opening hours of your business.
29. There are large and lovely trees in your local area.
30. The school reunion last night was a great success.
31. Pubs and bars dominate the city's social scene.
32. They had never ridden a bike in the snow before.
33. Dinner is now being served on a table in the garden.
34. Some people are very self-controlled and sure of themselves.
35. The art expert found a valuable painting being offered at a low price.
36. The child wore a white dress like her mother.
37. I was lying half asleep in the train station.
38. The girl started cleaning her house in preparation for the inspection.
39. Bushwalks are possible on both sides of the parks.
40. They sent us an order form from which we selected our items.
Mini-Dictation Sentences

Underlined words were missing on the test paper.

**Week 1 – Day 1**

1. The president was involved in the bombing incident.
2. I wanted to hire the band to give the place some atmosphere.
3. A boring novel is a great sleeping pill.
4. The group asked if I wanted to come along on the tour.
5. The government is planning a reform of the educational program.
6. The residents of the district were annoyed at the building plans.
7. My father was very fond of his dog, but my mother hated it.
8. We saw eight tiny drops of rain on the window.
9. Amongst her friends at university, she was considered beautiful.
10. The recent rainfall has caused very severe damage in the higher valleys.

**Week 1 – Day 2**

1. The reporter from the daily paper was responsible for the shocking story.
2. Down on the farm we used a cart for bringing food to the animals.
3. Laugh, dance and sing if fortune smiles upon you.
4. He threw several stones from the bridge while I quickly ran to hide.
5. The next local elections will take place during the winter.
6. The army officer was not happy about his posting to the desert.
7. The sports teacher made a point of riding her bike to work.
8. The fish began to leap wildly on the surface of the small lake.
9. It can be quite risky changing gear while half way up a hill.
10. The supermarket had to close due to economic problems.
**Week 1 – Day 3**

1. The manager of the farm will check on his bank account.
2. The policeman investigating the crime suggested that the door was forced open.
3. The children played games with beach balls for hours.
4. Was it the blue globe that broke when he switched on the light?
5. A hurricane was announced this afternoon on the TV.
6. The company director interviewed the basketball player.
7. In many country houses they still sound a bell for dinner.
8. Special task forces rescue hostages from kidnappers.
9. He turned around in an instant to look his attacker in the eye.
10. They didn’t hear the good news until last week on their visit to their friends.

**Week 1 – Day 4**

1. The old folk’s club had gone on its regular Friday afternoon bus trip.
2. The president’s wife wanted to hire a new cook before the summer.
3. The easygoing accountant relaxed throughout the voyage.
4. There’ll be big trouble if you touch that surface.
5. Artists have always been attracted by the life in the capital.
6. The victim of the gang’s vicious attack had cuts on his back.
7. The survey reports that university students are drinking more beer this year.
8. How much will it cost to do any necessary modernizing?
9. This new glove and shoe display attracts more customers than ever!
10. The local train left the station more than five minutes ago.
Week 2 – Day 1

1. Is a relaxed home atmosphere enough to help her reduce her fear?
2. He didn’t take the trip to Spain as he had planned.
3. The young boy got up quite early to watch the sun rise.
4. It would be difficult for the whole team to catch the bus to the mountains tonight.
5. Was she just naturally lazy about everything but her physical appearance?
6. In spite of technical progress, predicting the weather is still very difficult.
7. The history student decided to buy the second-hand car.
8. The ducklings in the pond were fighting for the breadcrumbs.
9. It’s unlikely that the date on which the battle occurred will be remembered.
10. The nurse looked after the fierce looking cut on John’s head.

Week 2 – Day 2

1. The demonstrators at the embassy were arrested by the riot police.
2. The art gallery in our street was opened only last week.
3. A toothpaste tube should be squeezed from the bottom.
4. Several changes have recently been made in the society’s regulations.
5. The bridge nearest the town was destroyed by terrorists.
6. I wouldn’t recommend having a big car in the city.
7. Shop assistants smile just long enough to make you feel important.
8. The flag ceremony affected me strongly and I was moved to tears.
9. Most of the scenes filmed on location in the hills were breathtaking.
10. His pet rabbit escaped from the garden after a visitor left the gate open.
Week 2 – Day 3

1. The writing on the walls of the underground station entertained the tourists.
2. The consumers’ association objected to the new delivery procedure.
3. I know I didn’t meet her early enough to change her mind.
4. The latest events have caused anger in the international community.
5. The road to the beach was washed out in the tropical rains.
6. Mary made a promise never again to laugh at tourists buying dolls.
7. My six cousins expected to be invited to stay all summer.
8. It seems as if Susan does all the floor cleaning for this household.
9. She was a very good policewoman, but she suffered in her off-duty hours.
10. Science has acquired an important place in Western society.

Week 2 – Day 4

1. That loud woman always comes up with strange examples.
2. The counting of the votes was interrupted when the computer failed.
3. He saw the badge of the traffic cop out of the corner of his eye.
4. In this famous coffee shop you will eat the best donuts in town.
5. Mothers usually leave the hospital two days after giving birth.
6. After wearing high heels my leg and back muscles feel tired.
7. Don’t you think her evening gown was a bit too formal for the occasion?
8. The art gallery had several paintings which turned out to be fakes.
9. The executive’s opinion was reported on the evening news.
10. The happy couple received many valuable wedding presents.
Week 3 – Day 1

1. There is not enough trained staff to meet the increasing demand for aged care services.
2. The young monkeys in the zoo often jumped on their keeper.
3. John realised that flying would save him a day in travel time.
4. The truth is, I wish I lived a little bit nearer to the supermarket.
5. It is getting much harder these days to find a place in a childcare centre.
6. The first concert given at the opera was an outstanding success.
7. Book clubs have been growing in popularity over many years.
8. In spite of technical progress, predicting the weather is still very difficult.
9. After troubled teenage years he went on to run an award winning restaurant.
10. The captain of the ship refused to turn back despite the illness of two team members.

Week 3 – Day 2

1. The crocodile began to swim towards the boat to be hand fed.
2. Big business is profiting from the increasing demand for pet insurance.
3. Green technology is being used to develop new race cars.
4. The politician was anxiously waiting for his chance to be leader.
5. Peter went to the party, but didn’t answer any questions.
6. A lot of people in the community don’t realise that the training is free.
7. The crowd lined the streets for hours in the hope of seeing the pop star.
8. As we walked toward the back of the room, we saw brightly coloured t-shirts.
9. The woman is working on a project she is hoping to finish next year.
10. We went to the public meeting to ask the government some questions.
Week 3 – Day 3

1. This is the first time an international exhibition has come to this town.
2. The baby monkey was born in a research centre and taken from its mother.
3. In this case the easier solution seems to appeal to the court.
4. After moving house four times the family finally found a place they really liked.
5. Unhealthy eating greatly increases your risk of serious disease or death.
6. The manager of the department store fired the lazy salesman.
7. The first flowers have bloomed because of August’s unusual warmth.
8. The director remembered the town where he was born and spent his youth.
9. This football season promises to be a very exciting one.
10. The medical association is considering recognizing internet addiction as a psychological condition.

Week 3 – Day 4

1. Did he like the smell of burning rubber and cars driving really fast?
2. Rain can bring the feeling of comfort and the sense of being safe.
3. The town hall was where the man had also married his first wife.
4. The angry people started throwing stones and set fire to two cars.
5. Health care workers want their conditions, pay and job security improved.
6. Having the occasional cigarette is not as harmless as you think.
7. Patient's beds were moved to the ground floor of the hospital.
8. It is a mystery how the female detective lost her millions.
9. The army was confident they could clean all the oil from the beaches.
10. The leader did not provide details about how the new plan would work.
Week 4 – Day 1

1. The artist created beautiful paintings and images for clothing.
2. I am the only one of my friends who doesn’t illegally copy movies.
3. The reporter wondered what was really happening inside the detention centre.
4. The effects of the disasters are still being dramatically felt.
5. We ride our bikes to get groceries, visit friends and go to the pub.
6. The government didn’t know what had happened at the demonstration.
7. The man changed his mind about asking for a divorce.
8. The short animated film is about the life of turtles.
9. Scientists have captured images of a planet forming around a star.
10. The teaching style of the workshop is gentle and focused with no technology used.

Week 4 – Day 2

1. This week he met an attractive older lady at an art auction.
2. Never take your laptop with you on a romantic weekend away.
3. Light streamed through the windows and created a sense of warmth.
4. I have always wondered why things taste bad to some people.
5. It’s important to be honest and truthful in your dealings with others.
6. She is feeling better and has started planning a birthday party.
7. The internet helps many social groups to form around a shared interest.
8. Please read the contract carefully so that you understand all the conditions.
9. How do birds move so quickly and safely through small spaces?
10. The little girl said many things that made her parents laugh.
Week 4 – Day 3

1. They say being in the country is all about living the quiet life.
2. Time spent at work can have a real impact on our health.
3. They saw the horses being delivered to the track on trucks.
4. The young people shared thoughts about living in a country town.
5. The art student tried to raise money for deaf people.
6. Photographing water is a great way to learn how your camera works.
7. Cute and curious animals were seen playing around the lake.
8. Hundreds of thousands of fans celebrated the team’s victory.
9. There is no cure for the condition, but diet can control the symptoms.
10. The boy volunteered with the local emergency services on his sixteenth birthday.

Week 4 – Day 4

1. Maybe you don’t care what you receive for Christmas this year.
2. Why not invest your wealth in changing the world?
3. There were more bugs in the house in the summer months.
4. They were walking down the street when the leaves were falling.
5. Space is always very important in the print media.
6. When he was gone I immediately turned to my friend in surprise.
7. You may be working from your car or home in the future.
8. The scientist denied lying to the government about his research.
9. Women worldwide have fought for the same voting rights as men.
10. The cost of getting a haircut is very expensive these days.
**Week 5 – Day 1**

1. Seaweed is moving south **to escape warming oceans and may risk becoming extinct.**
2. The driver screamed **when a huge spider ran across the window.**
3. She made money from **painting houses after the fires.**
4. Researchers have found that some **cancer cells respond better to treatment than others.**
5. Is looking at the computer screen **every day damaging your eyes?**
6. The smell of smoke was so strong that they found it hard to sleep.
7. Are you confused about the true facts of healthy eating?
8. The children weren’t excited by the thought of dressing up this year.
9. The boy lived at the hospital until he was about three years old.
10. The professor wasn’t very good at taking things apart and putting them back together.

**Week 5 – Day 2**

1. Why is the colour of **sand on some beaches different than on others?**
2. The couple carried **their letterbox with them every time they moved.**
3. One of my earliest memories was **sitting at this little table eating breakfast.**
4. The local bus service was **operated by a retired postman.**
5. In the last decade **computer games have grown in popularity throughout Italy.**
6. The workmen realised how much skill and time had gone into the project.
7. Low volume and poor reception on mobile phones make it hard to hear.
8. We are consuming more resources than the planet produces each year.
9. The current plan fails to address the region’s economic problems.
10. No-one really knows how many people there are on Earth.
Week 5 – Day 3

1. She continued on and crossed the finish line even though she felt sick.
2. Climate change is affecting the world’s oceans faster than scientists expected.
3. Watching movies under the stars has become a popular summer tradition.
4. There’s no doubt that technology can bring positives to our relationships.
5. The woman routinely bought a lot of things she didn’t need.
6. Each year millions of mobile phones are thrown away.
7. People still don’t accept that alcohol harms their health.
8. The new chairs became dirty after a few weeks of use.
9. Here you can eat food from all around the world in one weekend.
10. She was very happy not to have any more classes for the year.

Week 5 – Day 4

1. The man woke up in the garden with twenty-four mosquito bites.
2. Guide dogs are welcome throughout the building and at performances.
3. The most successful business people have a team of professionals around them.
4. Smokers should note that all restaurants and bars are now non-smoking.
5. Although he was shy, he was a very good communicator.
6. Bargaining is discouraged in shops, but may be an option at markets.
7. Some major fashion labels have their own retail outlets.
8. I was dying to tell her all about my plans.
9. On summer weekends there are free music performances in many parks.
10. He stood in the centre of the room and looked around with interest.
Experiment Three Results

See Chapter 7 for further discussion.

7.3.1 Dictations

7.3.1.1 Pre-test

An independent t test indicated that, while the mean scores of the W1C2 group were higher than those of the C1W2 group, the difference between the two groups was not statistically significant in either the partial dictation component of the pre-test, \( t(10) = 1.22, p = .25 \) (W1C2: \( M = 67.60, SD = 9.53 \); C1W2: \( M = 58.00, SD = 15.47 \)), or in the whole dictations, \( t(10) = -.58, p = .58 \) (W1C2: \( M = 53.20, SD = 8.84 \); C1W2: \( M = 49.43, SD = 12.42 \)).

Missegmentations

Independent t tests were conducted on the number of missegmentations in the pre-test. Again, there was no significant difference between the groups in the partial dictations, \( t(10) = .30, p = .77 \) (W1C2: \( M = 4.80, SD = 1.30 \); C1W2: \( M = 4.57, SD = 1.27 \)), nor the whole dictations, \( t(10) = -1.01, p = .33 \) (W1C2: \( M = 9.60, SD = 2.19 \); C1W2: \( M = 8.00, SD = 3.00 \)).

Correlations

There was no correlation between combined group performance in the dictation and missegmentation scores for the partial dictations \( r(10) = -.15, p = .65 \), nor whole dictations \( r(10) = .14, p = .67 \).

7.3.1.2 Pre-test – Post-test 1

In the partial dictation component, the main effect of ‘test number’ was not significant, \( F(1, 7) = .62, p = .46 \), nor the main effect of ‘training group’, \( F(1, 7) = 2.8, \)
$p = .15$. There was no significant ‘test number’ by ‘training group’ interaction, $F(1, 7) = .07, p = .80$.

In the whole dictation component, there was again no significant main effect of ‘test number’, $F(1, 7) = .05, p = .83$, or of ‘training group’, $F(1, 7) = 1.12, p = .33$, and no significant interaction between ‘test number’ and ‘training group’, $F(1, 7) = 1.85, p = .22$.

**Missegmentations**

An analysis on the number of missegmentations made in the partial dictation components of the pre-test and post-test 1, there was no significant main effect of ‘test number’, $F(1, 7) = .09, p = .78$, or of ‘training group’, $F(1, 7) = .58, p = .46$. There was also no significant interaction between ‘test number’ and ‘training group’, $F(1, 7) = 2.42, p = .17$.

The whole dictation analysis also shows no significant main effect of ‘test number’, $F(1, 7) = .14, p = .72$, or ‘training group’, $F(1, 7) = 2.28, p = .18$. There was also no significant interaction between ‘test number’ and ‘training group’, $F(1, 7) = .06, p = .82$.

**Correlations**

A strong, negative correlation between partial dictation scores and partial dictation missegmentation scores was significant, $r(8) = -.83, p < .01$, showing the higher the number of missegmented words, the lower the overall dictation score. There was no correlation between whole dictation and missegmentation scores, $r(8) = .01, p = .98$.

**7.3.1.3 Pre-test – Post-test 2**

The mean percentages of correct answers in the partial and whole dictation components of each test were analysed. In the partial dictations, the performance of
both training groups declined. There was a significant effect of ‘test number’, $F(1, 7) = 10.14, p = .02$. However, there was no significant effect of ‘training group’, $F(1, 7) = .11, p = .75$, nor interaction of ‘test number’ and ‘training group’, $F(1, 7) = .41, p = .55$.

For the whole dictation component the results are very different, with the mean scores of both training groups increasing in post-test 2. The main effect of ‘test number’ was significant, $F(1, 7) = 6.16, p = .048$. Again, there was no significant effect of ‘training group’, $F(1, 7) = .002, p = .96$, nor interaction between ‘test number’ and ‘training group’, $F(1, 7) = .82, p = .40$.

**Missegmentations**

The mean number of missegmentations made by each training group in the partial dictation component rose from the pre-test to post-test 2 and there was a significant main effect of ‘test number’, $F(1, 7) = 6.67, p = .04$. The main effect of ‘training group’, $F(1, 7) = .04, p = .84$, was not significant, nor was the ‘training group’ by ‘test number’ interaction, $F(1, 7) = .07, p = .80$.

In the whole dictation component there was no significant main effect of ‘test number’, $F(1, 7) = .02, p = .89$, ‘training group’, $F(1, 7) = 2.12, p = .20$, or interaction of ‘test number’ by ‘training group’, $F(1, 7) = 2.83, p = .14$.

**Correlations**

There was no significant correlation between partial dictation and partial missegmentation scores $r(8) = -.55, p = .16$, nor whole dictation and whole missegmentation scores, $r(8) = -.54, p = .16$. 
7.3.2 Mini-Dictations

7.3.2.1 Session 1

The main effect for ‘test number’ was significant, $F(5, 7) = 3.24, p = .02$. Pairwise comparisons showed a significant effect for ‘test number’ between ‘Test number 1’ and ‘Test number 4’, $p = 0.02$, and ‘Test number 1’ and ‘Test number 5’, $p = 0.05$, with the mean percentage correct being higher in Test 1 and dropping in the later tests.

The main effect of ‘training group’ was not significant, $F(5, 7) = 1.31, p = .30$, and nor was the ‘training group by test number’ interaction, $F(5, 7) = .96, p = .46$.

Session 1 Missegmentations

An analysis of the missegmentations made over the first six mini-dictations across the two training groups showed there was no significant main effect of ‘test number’, $F(5, 7) = .29, p = .91$, or ‘training group’ $F(5, 7) = .06, p = .82$, or interaction between ‘training group and test number’ $F(5, 7) = .03, p = .87$, on mean missegmentation scores.

Correlations

Pearson product-moment correlations were performed across the mini-dictation and missegmentation scores. In each case there was a negative correlation, but none reached statistical significance.

7.3.2.2 Session 2

An analysis was performed on the mini-dictation scores of session 2. The main effect of test number’ was significant, $F(5, 9) = 3.47, p = .01$. The main effect of ‘training group’ was not significant, $F(5, 9) = 2.19, p = .18$, and nor was the interaction between ‘training group and test number’ $F(5, 9) = 1.10, p = .38$. 

329
Although not significant, the ‘training group’ by ‘test number’ interaction was analysed by means of simple effect analyses performed separately for the two training groups. While the simple effect for the C1W2 group, which received word-spotting in this session, was not significant, the \( p \) value was low, \( F(5, 3) = 2.44, p = .07 \). This suggests greater statistical power, had there been larger participant numbers, may have produced a significant result. As expected, the W1C2 group by ‘test number’ effect was also not significant, \( F(5, 3) = 1.61, p = .20 \).

**Session 2 Missegmentations**

The mean number of missegmentations in the mini-dictations was analysed. The main effect of ‘test number’ was not significant, \( F(5, 9) = .83, p = .53 \), nor was the main effect of ‘training group’, \( F(5, 9) = .92, p = .37 \), nor the interaction of ‘training group’ and test number’, \( F(5, 9) = .58, p = .72 \).

**Correlations**

Pearson product-moment correlations showed a very strong negative correlation, \( r(10) = -.82, p = .004 \), between mini-dictation and missegmentation scores in the case of Test 3. No other analyses reached significance.

**7.3.3 Word-Spotting Scores**

An analysis was performed on the daily results achieved in the word-spotting training task across six training days. The two groups were combined for this analysis. The main effect of test number was significant, \( F(5, 8) = 9.89, p = <.001 \).

**7.3.3.1 Correlations**

Pearson product-moment correlations were performed on the score achieved on the word-spotting task and the number of missegmentations made in that day’s mini-dictation.
The correlation analysis of Day 1 showed a strong, negative relationship between word-spotting and missegmentation scores, which was significant, \( r(9) = -0.84, p = .004 \) (the higher the word-spotting score, the lower the number of missegmentations made in the mini-dictations). A negative correlation was also shown on the following days, but did not reach significance in any case: day 2, \( r(9) = -0.56, p = .10 \); day 3, \( r(9) = -0.37, p = .33 \); day 4, \( r(9) = -0.04, p = .93 \); day 5, \( r(9) = -0.19, p = .63 \). The analysis of Day 6 showed a positive correlation, but again this was not significant, \( r(9) = 0.22, p = .58 \).
Participant Questionnaire

We compared two types of training that were intended to help you listen to speech.

3. Which type was the MOST helpful?

Circle one: speech awareness word-spotting

Why?

____________________________________________________________________

____________________________________________________________________

4. Which type was the LEAST helpful?

Circle one: speech awareness word-spotting

Why?

____________________________________________________________________

____________________________________________________________________

How do you think this training could be improved?

____________________________________________________________________

____________________________________________________________________
This questionnaire refers to the **speech awareness training**.

Circle all the words that are true for you.

“The training was...”

too easy  easy  difficult  too difficult

fun  interesting  boring

helpful  confusing

too long  long  short  too short

tiring  exciting

Or add your own ideas: ____________________________________________

___________________________________________________________________
This questionnaire refers to the word-spotting training.

Circle all the words that are true for you.

“The training was...”

too easy    easy    difficult    too difficult

fun        interesting    boring

helpful    confusing

too long    long    short    too short

tiring    exciting

Or add your own ideas: __________________________________________

______________________________________________________________
New Vocabulary: 1

**macabre** strange and frightening, and often connected with death and injury.  
*She has a macabre fascination with true-life crime stories.*

**hospice** a place where people who are dying live and are cared for.  
The emphasis is on comfort, not curing.  
*The nurses at the hospice made the final weeks of John’s life comfortable.*

**aghast** filled with horror or shock.  
*When he told her the news, she looked at him aghast.*

**espouse** to support a belief or way of life; adopt; make one’s own; hold a belief.  
*The new theory has been espoused by many leading scientists.*

**placid** 1. A placid person is calm and does not often get angry or upset.  
*One son had a stormy nature. In contrast, the other son was more placid.*  
2. calm and peaceful  
*The lake was placid and still under the moonlight.*

**feisty** 1. (POS) active, confident, and determined:  
*The feisty young woman wouldn’t give up.*  
2. (NEG) easily offended or upset; repeatedly argues with people  
*He is difficult to work with. He gets feisty when anyone discusses politics.*

**pithy** A pithy remark expresses something in a very clear and direct way (brief, forceful and meaningful).  
*His letters, always pithy and forceful, appeared regularly in many newspapers.*

**usurp** to take someone else's job or power when you should not  
*Some people have accused the prince of trying to usurp the king's power.*

**bemused** slightly confused  
*I must admit that I was rather bemused by his sudden anger.*  
*her bemused expression*

**wean** 1. To start to give a baby food to eat instead of its mother's milk.  
*The kitten was weaned and fed by its owner with a bottle.*  
2. To get someone used to managing without something on which they have become dependent or of which they have become excessively fond.  
*She weaned herself from cigarettes.*
Choose the correct word for each sentence…

1. The couple was lucky that their first baby was very ____________ and calm.
2. Those ____________ unpopular views were often excluded.
3. The woman made a ____________ discovery when she was walking her dog.
4. The principal was ____________ at the violence she saw on the school’s playground.
5. I was saddened when my mother moved to the ____________, but I am confident that she is being made comfortable there.
6. Who is that ____________ girl who is always arguing with her teachers?
7. My sister ____________ her baby when he was three months old and started him on powdered milk.
8. There were a couple of attempts to ____________ the young king.
9. The actor looked slightly ____________ by the reporters’ questions.
10. The man welcomed his wife’s ____________ comments about politics.
New Vocabulary - 2

wisp n 1. a small thin line of cloud/smoke/steam
2. a thin piece of hair/grass, etc
   *He brushed a wisp of hair from her eyes.*

salvage v (from a place) to save things from a place where other things have been damaged or lost
   *gold coins salvaged from a shipwreck*
   (from a situation) to try and make a bad situation better
   *an attempt to salvage her reputation*

sachet n 1. a small bag containing a small amount of something
   *sachets of sugar and coffee powder*

pageant n 1. a show that happens outside in which people dress and act as if they are from a time in history
   *a traditional Christmas pageant; the Man from Snowy River pageant*
2. a beauty contest
   *the Miss World beauty pageant*

shove v to push someone or something in a rough way
   *He wouldn’t move, so I shoved him out of the way.*

turf n short thick grass and the soil it is growing in
   *This is a tough turf, especially suited for the harsh Australian climate.*

oblique adj not expressed in a direct way
   *an oblique comment*
   *I was angered by his oblique claim that the project had been his idea.*

lanky adj A lanky person is very tall and thin.
   *The lanky basketballer was surrounded by the other team.*

languid adj moving or speaking slowly and with little energy, often in an attractive way
   *a languid manner, voice*
   *She walked languidly towards him, her hips swaying.*

stifle v to stop something from happening or continuing
   *to stifle a sneeze/yawn*
   *Big supermarkets stifle competition*
Choose the correct word for each sentence….

1. The famous Carnival Queen promises to be the best so far with non-stop entertainment, fun, and thrilling performances.

2. Angry words lead Mary to Susan.

3. It was a beautiful day. There wasn’t a of cloud in the sky.

4. Just pour the chocolate mixture from your into a mug and stir in hot water.

5. The family tried to clothing and personal items from their home after the fire.

6. The dancer had a long, bony, body.

7. Golfers have been playing on artificial for years.

8. The current government policies the attempts of small businesses to be successful.

9. Life on the island is so unhurried and , you’re sure to have an absolutely relaxing holiday.

10. It is best not to directly challenge the boss, but to take an approach.
Find a synonym or antonym

<table>
<thead>
<tr>
<th></th>
<th>Synonym</th>
<th></th>
<th>Antonym</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>sachet</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>pageant</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>languid</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>salvage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>wisp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>turf</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>lanky</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>oblique</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>stifle</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>shove</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Total Score: ________________

Good luck!
### SLAM

<table>
<thead>
<tr>
<th>wisp</th>
<th>salvage</th>
</tr>
</thead>
<tbody>
<tr>
<td>sachet</td>
<td>pageant</td>
</tr>
<tr>
<td>shove</td>
<td>turf</td>
</tr>
<tr>
<td>oblique</td>
<td>lanky</td>
</tr>
<tr>
<td>languid</td>
<td>stifle</td>
</tr>
</tbody>
</table>
Experiment Four Results

See Chapter 8 for further details.

8.3.1 Dictations

8.3.1.1 Pre-test

An independent $t$ test found there was no statistically significant difference between the means of the set of partial dictations, $z(N = 31) p = .92$ or whole dictations, $t(29) = .56, p = .58$ for the two training groups.

Missegmentations

Independent $t$ tests were conducted on the number of missegmentations in each section of the pre-test. There was no significant difference between the two groups in the partial dictations, $t(28) = -.40, p = .69$, nor in the whole dictations, $t(29) = .49, p = .66$.

Correlations

A very strong, negative relationship between partial dictation scores and partial dictation missegmentation scores was significant, $r(31) = -.59, p < .001$, as was the very strong negative correlation between whole dictation scores and whole dictation missegmentation scores, $r(31) = -.61, p < .001$.

8.3.1.2 Pre-test – Post-test 1

The reported analyses are divided into partial and whole dictation results. The main effect of ‘test number’ in the partial dictations was significant, $F(1, 26) = 99.10, p < .001$. Partial dictation scores in post-test 1 were significantly higher than in the pre-test (pre-test: $M = 59.54, SD = 3.26$; post-test 1: $M = 73.99, SD = 2.77$). However, the main effect of ‘training group’ was not significant, $F(1, 26) = .37, p = .55$. There was no
significant interaction of ‘training group by test number’ $F(1, 26) = .18, p = .68$.

The main effect of test number was not significant in the analysis of the whole dictation scores, $F(1, 26) = .48, p = .50$, nor was the main effect of training group, $F(1, 26) = .22, p = .64$, nor the ‘training group by test number’ interaction, $F(1, 26) = .67, p = .42$.

**Missegmentations**

An analysis was conducted on the missegmentations made across the pre-test and post-test 1 for the 27 participants who attended 78% or more of the training. For the partial dictations the main effect of test number was significant, $F(1, 26) = 26.12, p = <.001$. The main effect of ‘training group’ was not significant, $F(1, 26) = .00, p = 1.00$. The ‘training group’ by ‘test number’ interaction was also not significant, $F(1, 26) = .25, p = .62$.

The analysis was repeated for the number of missegmentations made in the whole dictations in the pre-test and post-test 1. Here the main effect of ‘test number’ was not significant, $F(1, 26) = 1.78, p = .19$, and nor was the main effect of ‘training group’, $F(1, 26) = .11, p = .74$. The interaction of ‘training group’ by ‘test number’ was also not significant, $F(1, 26) = .002, p = .97$.

**Correlations**

Pearson product-moment correlations were performed between the post-test 1 dictation and missegmentation scores. As in the pre-test analysis, a very strong, negative relationship between partial dictation scores and partial dictation missegmentation scores was significant, $r(27) = -.57, p = < .001$. A negative correlation between whole dictation scores and whole dictation missegmentation scores almost reached significance, $r(27) = -.38, p = .051$.  

343
8.3.1.3 Pre-test – Post-test 2

There was a significant main effect of ‘test number’, \( F(1, 9) = 6.37, p = .04 \), with the performance of both training groups improving. The mean score of the groups combined was 57% in the pre-test (\(SD = 6.30\)) and 63% in post-test 2 (\(SD = 5.06\)). However, there was no significant main effect of ‘training group’ \( F(1, 9) = .01, p = .91 \). The ‘training group by test number’ interaction was insignificant, \( F(1, 9) = .64, p = .48 \).

An analysis of the whole dictations combined shows the main effect of ‘test number’ was not significant, \( F(1, 9) = .57, p = .47 \), and nor was the main effect of ‘training group’ \( F(1, 9) = .07, p = .80 \). The ‘training group by test number’ interaction was also not significant, \( F(1, 9) = .64, p = .45 \).

**Missegmentations**

An analysis of the mean number of missegmentations made in the pre-test and post-test 2 showed that in the partial dictation component, the main effect of ‘test number’ was not significant, \( F(1, 9) = 2.04, p = .19 \). There was also no significant main effect of ‘training group’, \( F(1, 9) = <0.001, p = 1.00 \), or interaction of ‘training group’ by ‘test number’, \( F(1, 9) = .004, p = .95 \).

In the whole dictation component, there was also no significant main effect of ‘test number’, \( F(1, 9) = 2.43, p = .16 \), or ‘training group’, \( F(1, 9) = .05, p = .82 \). There was also no significant interaction of ‘training group’ by ‘test number’, \( F(1, 9) = .05, p = .82 \).

**Correlations**

The negative relationship between the partial dictation scores and partial dictation missegmentation scores was not significant, \( r(10) = -.48, p = .16 \), and nor was the negative relationship between the whole dictation scores and whole dictation missegmentation scores, \( r(10) = -.17, p = .63 \).
8.3.2 Mini-Dictations

8.3.2.1 Partial Dictations

An analysis of the partial component results of the first six mini-dictations showed that the main effect of ‘test number’ was significant, $F(5, 21) = 20.51, p = < .001$. The main effect of ‘training group’ was not significant in the partial dictation analysis, $F(5, 21) = .67, p = .65$, and nor was the ‘training group by test number’ interaction, $F(5, 21) = .04, p = .84$.

Pairwise comparisons showed a significant effect for many partial dictation tests, as shown in Table O.1.

<table>
<thead>
<tr>
<th>Time Point 1</th>
<th>Direction</th>
<th>Time Point 2</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>$&gt;$</td>
<td>Test 2</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 1</td>
<td>$&gt;$</td>
<td>Test 5</td>
<td>$p = 0.004$</td>
</tr>
<tr>
<td>Test 1</td>
<td>$&gt;$</td>
<td>Test 6</td>
<td>$p = 0.01$</td>
</tr>
<tr>
<td>Test 2</td>
<td>$&lt;$</td>
<td>Test 3</td>
<td>$p = 0.02$</td>
</tr>
<tr>
<td>Test 2</td>
<td>$&lt;$</td>
<td>Test 4</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 3</td>
<td>$&lt;$</td>
<td>Test 4</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 4</td>
<td>$&gt;$</td>
<td>Test 5</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 4</td>
<td>$&gt;$</td>
<td>Test 6</td>
<td>$p = &lt; 0.001$</td>
</tr>
</tbody>
</table>

*Table O.1 Significant differences between the partial dictation component of the mini-dictation tests as shown by pairwise comparisons. Tests not mentioned were not significant.*

Missegmentations

The mean number of missegmentations made by the two training groups in the partial dictations of the first six mini-dictation tests was analysed. The main effect of ‘test number’ was significant, $F(5, 22) = 2.93, p = .02$. Pairwise comparisons showed the only significant difference was between Test 1 and Test 3, $p = .01$, with Test 1 having fewer missegmentations ($M = .93$) than Test 3 ($M = 2.25$). The results were due to a higher number of missegmentations across participants in both training groups and were not merely an artefact of one participant’s results.
The main effect of ‘training group’ was not significant, $F(5, 22) = 0.5, p = .82$, and nor was the interaction between ‘training group’ and ‘test number’, $F(5, 22) = 1.19, p = .32$.

**Correlations**

Pearson’s product-moment correlations ($r$) were performed between the dictation scores and missegmentation scores for each mini-dictation test. An analysis of the partial dictation component of the tests (with the one outlier removed), showed very strong significant negative correlations between the mean percentage of correct words in each dictation and the number of missegmentations made for all but two tests. In addition, Test 3 was close to reaching significance at the 0.05 level, $p = .066$. See Table 8.8 for the figures on correlations between the percentage of correct words and the number of words missegmented in the partial dictation component of the daily mini-dictation tests.

**8.3.2.2 Whole Dictations**

The mean scores of the two training groups over the first six mini-dictation tests were very similar. The main effect of ‘test number’ was significant, $F(5, 22) = 20.26, p = < 0.001$. The main effect of ‘training group’ was not significant, $F(5, 22) = .02, p = .88$. The ‘test number’ by ‘training group’ interaction was also not significant, $F(5, 22) = .11, p = .75$. Table O.2 shows the significant pairwise comparisons.

<table>
<thead>
<tr>
<th>Time Point 1</th>
<th>Direction</th>
<th>Time Point 2</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>&gt;</td>
<td>Test 6</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 2</td>
<td>&lt;</td>
<td>Test 4</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 2</td>
<td>&gt;</td>
<td>Test 6</td>
<td>$p = .01$</td>
</tr>
<tr>
<td>Test 3</td>
<td>&lt;</td>
<td>Test 4</td>
<td>$p = .02$</td>
</tr>
<tr>
<td>Test 3</td>
<td>&gt;</td>
<td>Test 6</td>
<td>$p = &lt; 0.001$</td>
</tr>
<tr>
<td>Test 4</td>
<td>&gt;</td>
<td>Test 6</td>
<td>$p = &lt; 0.001$</td>
</tr>
</tbody>
</table>

*Table O.2 Significant differences between whole-dictation tests as shown by pairwise comparisons.*
Missegmentations

The mean number of missegmentations made by both training groups in the whole dictation component of the daily mini-dictations was also analysed. The main effect of 'test number' was significant with a Huynh-Feldt adjustment to the degrees of freedom, $F(4.06, 85.33) = 12.59, \ p = < 0.001$.

The main effect of 'training group' was not significant, $F(5, 22) = 1.41, \ p = .25$.

However, the 'training group' by 'test number' interaction was significant, $F(4.06, 85.33) = 2.60, \ p = .04$. This interaction was analysed by means of simple effect analyses, which were performed separately for the two training groups. Using the Huynh-Feldt correction, the simple effect for Group W1C2 was significant, $F(4.32, 38.91) = 6.60, \ p = < 0.001$. The simple effect for Group C1W2 was also significant, $F(3.65, 43.84) = 7.96, \ p = < 0.001$.

Each simple effect analysis was followed by simple contrasts comparing the second to sixth tests to the first test. The simple comparisons showed that Group W1C2, which received word-spotting training, made significantly more missegmentations in Test 6 than they had in Test 1, $F(1, 9) = 6.98, \ p = .03$. This was led by one participant who made 12 missegmentations in Test 6, but only one in Test 1.

Group C1W2, which received training in connected speech processes in session 1, made significantly more missegmentations in Test 5, $F(1, 12) = 14.90, \ p = .002$, and Test 6, $F(1, 12) = 6.94, \ p = .02$, than they had in Test 1. However, they also made significantly less missegmentations in Test 4 than they had in Test 1, $F(1, 12) = 6.31, \ p = .03$, (Test 1 $M = 1.69$, Test 4 $M = 0.85$).
Correlations

Pearson’s product-moment correlation analyses on the whole dictation component of the tests produced only one significant negative correlation between the dictation scores and missegmentation scores, Test 1, \( p = .001 \). Analyses of all other tests showed negative correlations that were not significant.

8.3.4 Word-spotting Scores

8.3.4.1 Mean Percentage Correct

The two groups were initially combined and an analysis performed on the daily results achieved in the word-spotting training task. In the combined groups analysis, the main effect of ‘test number’ was significant, \( F(4, 20) = 9.22, p < .001 \). A comparison of the performance of the two training groups reveals a significant difference in the performance of the two groups. The mean effect of ‘test number’ was significant, \( F(4, 20) = 8.59, p < .001 \), as was the interaction between ‘test number’ and ‘training group’, \( F(4, 20) = 2.71, p = .036 \). The mean effect of ‘training group’ was not significant, \( F(4, 20) = .03, p = .86 \).

The ‘test number’ by ‘training group’ interaction was analysed by means of simple effect analyses performed separately for the two training groups using separate error terms. Each simple effect analysis was followed by simple contrasts comparing each level of test number to the pretest. The simple effect for Group W1C2 was significant, \( F(1, 7) = 4.44, p = .007 \), with the simple comparisons showing that the group’s mean score on Day 4 was significantly higher than it was on Day 1, \( F(1, 7) = 7.81, p = .027 \).

The simple effect for Group C1W2 was also significant, \( F(1, 12) = 8.19, p < .001 \), with the simple comparisons showing a significantly lower mean score on Day 3 than had been gained on Day 1, \( F(1, 12) = 6.86, p = .022 \).
8.3.4.2 Mean Percentage Segmented Correctly

An analysis was conducted on the segmentation data from the first five days of the training for the same participants as in the previous analyses. An analysis of the training groups combined resulted in a significant main effect of ‘test number’, $F(4, 20) = 19.37, p < .001$.

An analysis was performed between the training groups to detect any differences in performance. The main effect of ‘test number’ was significant, $F(4, 20) = 17.28, p < .001$. However, there was no significant main effect of ‘training group’, $F(4, 20) = .08, p = .78$, nor ‘training group’ by ‘test number’ interaction, $F(4, 20) = 1.21, p = .31$.

**Correlations**

Again, there was a strong significant correlation between the word-spotting and missegmentation scores on Day 1, $r(10) = - .77, p = .010$. There were negative correlations between the data from all further days, but none reached significance.

8.3.5 Reaction Time Post-test

An independent $t$ test was conducted on the mean scores for each participant in the reaction time post-test of session 1. The mean reaction time was lower in Group W1C2, which had received word-spotting training before the vocabulary teaching and RT test ($M = 1.78$ s, $SD = .59$ s), than Group C1W2, which had received training in connected speech processes ($M = 2.11$ s, $SD = 1.02$ s). Despite the results not reaching significance, $t(22) = .90, p = .38$, this outcome is promising.
Sound System Awareness Training

Syllable Stress

One of the most important features of spoken English is syllable stress. English has stressed and unstressed syllables, and unstressed syllables are often reduced. This means that the vowel sounds like *uh* or *ih* (usually *uh*). Dictionaries and pronunciation books usually use the symbol /ə/ for *uh*.

In each of the following words, the first syllable is stressed. It is a strong syllable. The second syllable is unstressed. It is a weak syllable.

**STRONG-weak (Oo)**

- buckle “BU-kuhl”
- writing “WRI-ting”
- milky “MIL-ki”
- number “NUM-buh”

The words below are stressed on the second syllable. The pattern is weak-strong.

**weak-STRONG (oO)**

- away “uh-WAY”
- prevent “pruh-VENT”
- repeat “ruh-PEAT”
- again “uh-GAIN”

It is useful to find a way to represent stressed/unstressed when you learn a new word. Some people use a stress mark (diacritic), e.g. ’doctor, pre’vent. In this lesson we will use the symbols Oo and oO, where O is stressed and o is unstressed, e.g. doctor = Oo, prevent = oO.

---

**Activity 1**

- a) Type your participant code in the ‘subject’ field.
- b) Click on browse.
- c) Choose file 01.rtf.
- d) Click on run.

If the word you hear is **STRONG-weak (Oo)**, press the left-shift key and the spacebar to continue.

If the word you hear is **weak-STRONG (oO)**, press the right-shift key and the spacebar to continue.

*Remember, for all activities:*

*When you see the screen that says ‘The End’, press ESC.*

*Please save the file at the end of each activity.*

*If you would like to do an activity again, click run.*
Function Words

Some words in spoken sentences are also unstressed. The more important words (containing the meaning needed to understand the sentence) are stressed and the less important words (containing the grammar of the sentence) are unstressed. These ‘grammar’ words are called function words. They are mainly:

- articles (the, a, an ...)
- prepositions (for, to, on, from, in ...)
- pronouns (he, she, it, us, them ...)
- conjunctions (and, but, so ...)
- auxiliary verbs (am, can, had, was ...)

The more important words in a sentence are called content words. They are:

- nouns (dog, chair, beauty, decision ...)
- verbs (run, speak, consider ...)
- adjectives (handsome, cold, deep ...)
- adverbs (quickly, carefully, often, last ...)

In a sentence, short function words will be reduced to the ‘uh’ sound.

a) Do they think so?  
   *duh they* (It sounds like d'they?)

b) I hear you.  
   *yuh*

c) Do you know?  
   *duh ya (d'ya)*

d) I know for sure.  
   *fuh (f'sure)*

e) Talk to me.  
   *tuh (t'me)*

Now compare the sentences,

a) He was away for months.  
   *fuh (f'months)*  
   *For* is a function word. It is reduced.

b) He was away four months.  
   *four (four months)*  
   *Four* is a content word. It is not reduced.
Activity 2
   a) Type your participant code in the ‘subject’ field.
   b) Click on browse.
   c) Choose file 02.rtf.
   d) Click on run.

Write the words you hear in the spaces below. There may be more than one word in each space.

1. I thought we would watch ________ movie ________ video.
2. We are having ________ neighbours ________ dinner.
3. I want ________ sleep ________ long ________ possible.
4. I’d like ________ go hiking or camp ________ beach.
5. I need ________ catch ________ on ________ work I didn’t get done this week.
6. I have ________ list a mile long ________ things ________ do!
7. I’ve ________ got to get ready ________ a math exam.
8. There’s ________ new exhibit ________ art museum I’d like to see.
Reductions for ‘and/or’

Pay attention to the reductions for ‘and’ and ‘or’. Because the sounds are so short, they might be confused or not heard at all. Short, reduced words can seem to disappear in spoken English.

a) Do you drink coffee or tea? coffee ‘uh’ tea
b) Let’s stop and eat lunch here. stop ‘n’ eat
c) Shall we walk or drive? walk ‘uh’ drive
d) Come and look. come ‘n’ look

Activity 3

a) Type your participant code in the ‘subject’ field.
b) Click on browse.
c) Choose file 03.rtf.
d) Click on run.

Choose the word you hear.

If you hear ‘and’, press the left-shift key and the spacebar to continue.

If you hear ‘or’, press the right-shift key and the spacebar to continue.

Press the spacebar to begin.

Activity 4

a) Look at the computer screen.
b) Click on browse.
c) Choose file 04.rtf.
d) Click on run.

Write the words you hear in the spaces below.

At the Store

A: Will that be cash __________ charge?
B: Cash. Wait, I mean cheque.

A: Do you have a driver’s licence __________ other identification?
B: How about a driver’s licence __________ passport?
A: One __________ the other is OK.
B: Here you are.
Activity 5

a) Type your participant code in the ‘subject’ field.
b) Click on browse.
c) Choose file 05.rtf.
d) Click on run.

Write the words you hear in the spaces below. There may be more than one word in each space.

Food Talk

1. A: __________________ another helping?
   B: I couldn’t eat another bite.

2. A: I’d like a hamburger _________ chocolate shake.
   B: _______ here ______ to go?

3. A: What do you ________________ order ___________ dessert?
   B: Let’s ask ________ see ________ dessert menu.

4. A: Where did you learn ___________ cook ________________?
   B: I grew _______________ family of cooks.

5. A: Why aren’t you eating ___________ vegetables? I thought
   ____________ favourite.
   B: They taste ______________ different.

6. A: Let’s stop _______________ bakery _______________ pastries.
   B: I don’t think there’s enough time.
Sound System Awareness Training

SPECIAL SOUND CHANGES

The Flap

When \( t \) and \( d \) occur between vowels, they are often pronounced as **flaps**. This sound is similar to a \( d \), but it is much faster. The tongue tip touches the tooth ridge very quickly. You may have heard native English speakers saying something that sounds like “waadda” instead of ‘water’ and “bedda” instead of ‘better’.

You may often see ‘got to’ written as ‘gotta’, but hear it said as ‘goda’. Here we see three processes taking place.

1. Because ‘to’ is a function word, its vowel is reduced to schwa, the \( uh \) sound.
   
   \[
   \text{got to} \rightarrow \text{got tuh}
   \]

2. Because the two ‘t’s are next to each other, they are reduced to one \( t \).
   
   \[
   \text{got tuh} \rightarrow \text{gotuh}
   \]

3. Because the \( t \) occurs between two vowels (even though those vowels are in different words), the \( t \) is said as a flap.
   
   \[
   \text{gota} \rightarrow \text{goda}
   \]

In the same way, you will often hear single words forming a flap when linked in a sentence as in,

a) What else do you want? \( \rightarrow \) “Whadelse do you want?”

b) I don’t know how to do it. \( \rightarrow \) “I don’t know howduhdo it”
   (Note the ‘uh’ sound in the function word, ‘to’.)

Between a vowel and \( l \) or \( r \), the \( t \) may be flapped and sound like \( d \).

a) title \( \rightarrow \) “tidle”

b) Her head hurt a lot. \( \rightarrow \) “Her head hurda lot.”
Activity 1

e) Type your participant code in the 'subject' field.
f) Click on browse.
g) Choose file 01.rtf.
h) Click on run.

Write the words you hear in the spaces below. There may be more than one word in each space.

At the Gym

1. A: Have you exercised enough ________________________________?
   B: No. I'm ______________________________ tired.

At School

2. A: I need to ____________________ bookstore ________________.
   B: I'll go with you.

In the Kitchen

3. A: Do you want the ________________________________?
   B: ____________________ good.

At a Restaurant

4. A: Why isn’t anyone ____________________ us?
   B: I'll go ____________________ find a waiter.

Remember, for all activities:
Please save the file at the end of each activity.
If you would like to do an activity again, click run.

Activity 2

e) Type your participant code in the 'subject' field.
f) Click on browse.
g) Choose file 02.rtf.
h) Click on run.

“What are you” and “What do you” are both pronounced whaddaya. The d is a flap. Use the grammar of the sentence to help you tell the difference. Listen to these sentences. Is the intended question “what are you” or “what do you”? Choose the correct answer.

If the sentence is “What are you”, press the left-shift key and the spacebar to continue.

If the sentence is “What do you”, press the right-shift key and the spacebar to continue.
Activity 3
   a) Type your participant code in the ‘subject’ field.
   b) Click on browse.
   c) Choose file 03.rtf.
   d) Click on run.

Write the words you hear in the spaces below. There may be more than one word in each space.

Two Friends
1. A: __________________________ you see what I’m giving you for your birthday!
   B: __________________________! My birthday is six months away.
2. A: I love your jacket.
   B: Thanks. I _________________________ on sale.
   A: Where did you __________________________ ?
   B: __________________________ little shop down the street.
   A: I __________________________ new coat.
   B: You __________________________ go there. I’ll show you where ____________.

-nt Reduction

When nt occurs in a word, many speakers leave out the t and use a flapped n.

   a) twenty     →     twenty
   b) identify    →     idenify
   c) wanted      →     wandid
   d) sentences   →     senences
Activity 4

a) Type your participant code in the ‘subject’ field.
b) Click on browse.
c) Choose file 04.rtf.
d) Click on run.

Put a tick (✓) next to the picture that best fits the sentence you hear.
Sound System Awareness Training

**SPECIAL SOUND CHANGES**

The Glottal Stop

Many native English speakers use a special sound to replace *t*. It is called a **glottal stop**. A glottal stop is formed when the vocal cords close completely for a moment, stopping the air. It is the sound you hear in the negative “uh uh”! It most often occurs when *t* is followed by a syllabic *n*. For example the sound of the word *eaten* can be divided into two syllables, *eat* and *n → eat-n*. If the *t* is replaced by a glottal stop, it sounds more like *ea’n*.

You may hear a glottal stop in these cases.

a) I’m **waiting** for the bus.  →  *wai’n*
b) Who hasn’t **eaten** yet?  →  *ea’n*
c) It’s **getting** late.  →  *ge’n*

A glottal stop can also replace a syllable as in,

a) That’s not **important**.  →  *impor’n*
b) Is something the matter?  →  *su’m*
Activity 1  
  a) Type your participant code in the ‘subject’ field.  
  b) Click on browse.  
  c) Choose file 01.rtf.  
  d) Click on run.  

Write the words you hear in the spaces below. There may be more than one word in each space.  

At Work  
  5.  A: There’s a meeting scheduled at 11:00  
      B: I’m sorry, but I’ve got ________________________ else planned.  

At a Cafe  
  6.  A: Sorry I’m late. What’s ________________________?  
      B: Not much. We’re just ________________________ for someone to take our order.  

At the Hairdressers  
  7.  A: Why don’t we ________________________ your hair?  
      B: Oh, no. I want to keep it natural.  

At School  
  8.  A: That food in the cafeteria made me thirsty. I need to find a drinking ________________________.  
      B: I’m really thirsty, too. I wish we hadn’t ________________________ such salty food.  

Of  

“Of” is a function word. Because it is a function word, it is commonly reduced to a.  

a) I’m sick of work.  → sicka  
b) It’s kind of difficult.  → kinda  
c) I want some of that.  → somea  

In the following examples, the sound a is preceded by a flap.  

a) She has a lot of money.  → lodda  
b) He’s out of time.  → oudda
Activity 2
   a) Type your participant code in the ‘subject’ field.
   b) Click on browse.
   c) Choose file 02.rtf.
   d) Click on run.

Write the words you hear in the spaces below. There may be more than one word in each space.

“Where have you been?”

1. I’ve been ______________________________ for a few days.
2. At the gym. I’m so ______________________________.
3. Traveling ______________________________ on business.
4. Looking for a job. I’ve been ______________________________ for a week.

Activity 3
   a) Look at the computer screen.
   b) Click on browse.
   c) Choose file 03.rtf.
   d) Click on run.

Write the words you hear in the spaces below. There may be more than one word in each space.

At a Restaurant

1. A: This dessert tastes ______________________________ world.
   B: Let me try a ______________________________ that.
2. A: Oh no! I’m ______________________________.
   B: How are we going to pay for ______________________________ food?
3. A: I’m really glad we came here tonight. I’m
   so ______________________________ cooking.
   B: But if we keep eating food like this, we’re going to be
   ______________________________ shape real soon.
Linking of with Vowels

When “of” is linked to a vowel, the sound is uv.

a) All of us went → alluvus
b) What do you think of her? → thinkuver

Sometimes two different pronunciations are common.

It’s out of order. → outuvorder
→ outaorder

Activity 4

a) Type your participant code in the ‘subject’ field.
b) Click on browse.
c) Choose file 04.rtf.
d) Click on run.

It is easy to confuse “of” with the article “a”. Listen to the sentences.
If you hear “a”, press the left-shift key and the spacebar to continue.
If you hear “of”, press the right-shift key and the spacebar to continue.
Activity 5
   a) Type your participant code in the ‘subject’ field.
   b) Click on browse.
   c) Choose file 05.rtf.
   d) Click on run.

Listen to each sentence. Does ‘of’ sound like “a” or “uv”?

If you hear “a”, press the left-shift key and the spacebar to continue.

If you hear “uv”, press the right-shift key and the spacebar to continue.

Activity 6
   a) Type your participant code in the ‘subject’ field.
   b) Click on browse.
   c) Choose file 06.rtf.
   d) Click on run.

Listen to the conversation. Write the words you hear in the spaces below. There may be more than one word in each space.

Two Friends at Work

A: What are you doing for lunch?

B: I thought I’d go run around the track. Would you like to join me?

A: To be honest, I’m ____________ shape. After just a few minutes, I’d be ____________ breath.

B: Well, that’s the best reason to go running. So you can get better.

A: Then I need to find some other _______________exercise. I’ve never really been fond ____________running.
Experiment Five Results

See Chapter 9 for further details.

### 9.3.1 ALL post-test

A 2 X 2 X 2 (training received X language background X ALL exposure) between-groups ANOVA was performed on the ALL test phase scores. There were no significant main effects of training, $F(1, 47) = .03, p = .86$, language, $F(1, 47) = 2.99, p = .09$, or ALL exposure, $F(1, 47) = 1.53, p = .22$. The ‘training by language’ interaction was also not significant, $F(1, 47) = .91, p = .35$.

The ‘language by ALL exposure’ interaction was significant, $F(1, 47) = 5.27, p = .027$, but not the ‘training by ALL exposure’ interaction, $F(1, 47) = 3.50, p = .07$. The interaction of ‘training by language by ALL exposure’ was also not significant, $F(1, 47) = .28, p = .60$.

Further analyses of the ‘language by ALL exposure’ interaction showed that the simple effect of language background was significant for ALL exposure set A, $F(1, 23) = 7.97, p = .007$, but not for ALL exposure set B, $F(1, 23) = .16, p = .69$. In addition, the simple effect of ALL exposure set was significant for the L1 group, $F(1, 23) = 6.13, p = .017$, but not for the L2 group, $F(1, 23) = .55, p = .46$.

There was no correlation between the length of residence (LOR) of the L2 group participants and their ALL task scores, $r(24) = -.13, p = .56$. There was also no correlation between the first language of the participants in the L2 group and their ALL test score, $r(24) = .10, p = .65$. 

364
9.3.2 Word-spotting Scores

9.3.2.1 Mean Percentage Correct

The mean percentage of correct word-spotting responses by the two language groups (L1 and L2) across the three daily training sets were analysed. There was a significant main effect for 'language', $F(2, 23) = 95.61, p = < 0.001$, with the L1 group receiving higher scores than the L2 group across all three training sets. The main effect of 'training set' was not significant, $F(2, 23) = .91, p = .41$, nor was the 'training set by language' interaction, $F(2, 23) = 2.14, p = .13$.

9.3.2.2 Mean Percentage Segmented Correctly

This time the main effect of 'training set' was significant (shown here with a Greenhouse-Geisser adjustment to the degrees of freedom following a violation of Mauchly's Test of Sphericity), $F(1.22, 26.82) = 5.64, p = .019$. Pairwise comparisons show that Set A03 was segmented significantly better than Set A01, $p = .024$.

The main effect of 'language' was also significant, $F(2, 23) = 54.01, p = <0.001$, with mean scores for the L1 group 10.44% higher than those of the L2 group. However, the interaction of 'training set by language' was not significant, $F(1.22, 26.82) = 1.81, p = .19$.

9.3.3 Correlations

There were positive correlations between the participants’ ALL task scores and their ability to segment the word-spotting material, but these failed to reach significance in any of the training sets: Set A01, $r(24) = .25, p = .24$, Set A02, $r(24) = .14, p = .53$, and Set A03, $r(24) = .13, p = .55$. 

365