Perceptual learning of context-sensitive phonetic detail

Katharine Joan Barden
Darwin College

March 2011

This dissertation is submitted for the degree of Doctor of Philosophy at the University of Cambridge

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text.

This dissertation does not exceed the 80 000 word limit set by the Modern and Medieval Languages Degree Committee.
Word count: 56 723 (excluding bibliography)

ACKNOWLEDGEMENTS

I would like to thank Professor Sarah Hawkins for her support, enthusiasm and inspiration for the past four years.

Thanks also to friends and colleagues, past and present, from the phonetics lab for stimulating discussions, consenting to be guinea pigs, and helping to make work fun! Special thanks are due to Naomi Hilton, Sarah Hawkins and Mark Jones for lending their voices, and to Geoff Potter for his patience and technical expertise.

And, of course, thank you to David for warning me in advance, and supporting me throughout.

This research was funded by a grant from the AHRC.

CONTENTS

Note to Examiners 1

Abstract 2

Chapter 1. Introduction

1.1 Overview 3

1.2 Chapter outline 4

1.3 Multiple levels of representation 4

1.3.1 Phonetic detail conditioned by local and long-domain phonetic context..... 5

1.3.2 Phonetic detail conditioned by syllabic context 6

1.3.3 Phonetic detail conditioned by higher levels of prosodic structure 7

1.3.4 Phonetic detail relating to grammatical function and meaning 7

1.3.5 Phonetic detail conditioned by interactional context 8

1.3.5.1 Turn-taking cues 8

1.3.5.2 Phonetic detail reflecting pragmatic and emotional meaning 9

1.3.5.3 Phonetic detail reflecting talker-related contexts 10

1.3.6 Multiple levels of representation and models of speech perception 11

1.4 What categories can become associated with phonetic detail during perceptual learning? 13

1.4.1 Evidence for perceptual learning at multiple levels of representation..... 13

1.4.2 Perceptual learning at sub-lexical and sub-segmental levels of structure 15

1.4.3 Phonetic detail conditioned by adjacent phonetic context 17

1.4.4 Phonetic detail conditioned by syllabic structure..... 20

1.4.5 Phonetic detail associated with phrase boundaries 22

1.4.6 Phonetic detail associated with pragmatic context 23

1.4.7 Cross-modal variation: Phonetic detail associated with visual context 24

1.4.8 'Talker-specific' perceptual learning 25

1.4.9 Perceptual learning of sociophonetic variation 27

1.4.10 Context-sensitive representations: Summary of evidence from perceptual learning 28

1.5 Theoretical motivation 29

1.5.1 Elucidating multiple levels of representation in an exemplar approach 29

1.5.2 Constraints on learnability 30

1.6 Outline of thesis 31

Chapter 2. Experiment 1: Perceptual learning of a stress-conditioned allophone

2.1 Background 32

2.1.1 Perceptual learning: The importance of natural stimuli and tasks 32

2.1.1.1 Synthetic speech ……..…….. 33
2.1.1.2 Encoding of relevant phonetic detail: The role of variability …… 34
2.1.1.3 Encoding of relevant phonetic detail: The role of attention ….…… 36
2.1.1.4 The importance of the assessment task …… 37
2.1.1.5 Summary of the naturalistic approach ……..... 38
2.1.2 Methodological background to Experiment 1 ……. 38
2.2 Method …… 39
2.2.1 Subjects …… 39
2.2.2 Materials …….. 40
2.2.2.1 Training materials: Story ……. 40
2.2.2.2 Test materials: Auditory lexical decisions …… 43
2.2.3 Procedure ……..... 44
2.3 Hypotheses …….. 45
2.4 Results and Discussion ……. 46
2.4.1 Data pre-processing …… 46
2.4.2 Lexical decision task: Lexical acceptance rates …… 47
2.4.2.1 Statistical technique for binary responses …… 47
2.4.2.2 Analyses of lexical acceptance rates …… 47
2.4.3 Discussion of lexical acceptance rates …… 49
2.4.3.1 Stressed-vowel stimuli ……. 49
2.4.3.2 Unstressed-vowel stimuli …… 51
2.4.4 Lexical decision task: Response times …….. 52
2.4.5 Discussion of response time analyses …… 53
2.5 Conclusion ……..... 55

Chapter 3. Experiments 2a and 2b: Perceptual learning of a duration difference associated with grammatical category

3.1 Background …….. 58
3.1.1 Experiment 2a: Theoretical background – Listeners' sensitivity to phonetic detail reflecting grammatical category …….. 58
3.1.2 Experiment 2a: Methodological background ……..…… 62
3.1.2.1 Selection of duration as the cue to be manipulated …….. 62
3.1.2.2 Test task.…… 62
3.2 Experiment 2a: Method ……... 63
3.2.1 Subjects ……..…… 63
3.2.2 Materials ……..…… 63
3.2.2.1 Training materials: Story …….……. 63
3.2.2.2 Test materials: Sentence completion ……... 65

3.2.3 Procedure …… 67
3.3 Experiment2a: Hypothesis ……..… 68
3.4 Experiment 2a: Results …… 68
3.4.1 Data pre-processing ……. 68
3.4.2 Analysis of verb and noun responses …….. 68
3.5 Experiment 2a: Discussion …….. 70
3.6 Experiment 2b: Methodological background …….. 72
3.7 Experiment 2b: Method ……..…….. 72
3.7.1 Subjects ……. 72
3.7.2 Materials …….. 73
3.7.2.1 Training materials: Story ….……. 73
3.7.2.2 Test materials: Sentence completion … ……... 73
3.7.3 Procedure ……. 74
3.8 Experiment 2b: Hypothesis ……. 75
3.9 Experiment 2b: Results …… 75
3.9.1 Data pre-processing ……. 75
3.9.2 Analysis of verb and noun responses …….. 75
3.10 Experiment 2b: Discussion …… 76
3.10.1 Differences between Experiment 2a and Experiment 2b …….. 76
3.10.1.1 Procedural difference ……. 76
3.10.1.2 Differences in experimental sentence beginnings …….. 76
3.10.2 Effect of Task number ….……. 79
3.11 General discussion of Experiments 2a and 2b …….…….. 80
3.11.1 Why was there no learning effect for Short words in Experiment 2a? …. 80
3.11.2 A role for prior associations in constraining perceptual learning ……. 80
3.11.3 What was learned in Experiment 2a? …… 81
3.12 Conclusion ……..…… 82

Chapter 4. Experiment 3: Perceptual learning of phonetic detail that indicates morphological structure

4.1 Background …….……. 83
4.1.1 Phonetic detail and morphology ……. 83
4.1.2 Methodological background …… 85
4.2 Method ……..…… 87
4.2.1 Subjects ……. 87
4.2.2 Materials ……..……... 88
4.2.2.1 Training materials: Stories…… 88
4.2.2.2 Test materials: Intelligibility-in-noise sentences ……. 89

4.2.3 Procedure 91
4.3 Hypothesis 92
4.4 Results 92
4.4.1 Data pre-processing 92
4.4.2 Analysis of intelligibility-in-noise task 92
4.4.2.1 Keyword correct 92
4.4.2.2 Keyword of suffixed variant correct 95
4.4.2.3 Morphological type correct 95
4.4.3 Filler sentences in intelligibility-in-noise task 95
4.4.3.1 Target word correct 96
4.5 Discussion 97
4.5.1 Perceptual learning by Accent subjects is sensitive to morphological context 97
4.5.1.1 Differences in scoring methods 98
4.5.2 Co-existence of context-sensitivity and generalisation 99
4.5.2.1 Generalisation due to adaptation of category representations at multiple levels of structure 100
4.5.2.2 Generalisation due to task-level responses 101
4.5.3 The influence of Morphological type on perceptual learning by Control subjects 102
4.6 Conclusion 104

Chapter 5. General Discussion

5.1 Introduction 106
5.2 Proposed properties of a model of perceptual learning 106
5.3 Adaptation to phonetic associations at multiple levels of representation 108
5.4 Factors influencing the learnability of an association 113
5.4.1 Activation of phonetic properties 113
5.4.2 Activation of prior categories 114
5.4.2.1 Category 'robustness' 115
5.4.2.2 Bottom-up activation of categories because of prior associations .. 116
5.4.2.3 Other factors affecting activation of previously-learned representations 117
5.4.3 How do the factors that affect learnability interact? 117
5.4.4 The stability-plasticity relationship: The role of prior knowledge 118
5.5 Methodological implications of the present research 119
5.5.1 Multiple levels of representation..... 119
5.5.2 Generalisability of conclusions 120

5.5.3 Minimising or monitoring perceptual learning during assessment…… 122
5.5.4 How perceptual learning effects are manifested …….. 122
5.6 Conclusion ……. 123

Bibliography ……... 124

Appendices

Appendix A ……... 139
Appendix B …….. 148
Appendix C ……...…… 149
Appendix D …….. 150
Appendix E ……... 151
Appendix F …… 155
Appendix G ……... 157
Appendix H ……... 158
Appendix I ……. 159
Appendix J ……160
Appendix K ……162
Appendix L …… 169
Appendix M …….. 170
Appendix N ……... 172

NOTE TO EXAMINERS

Sample sound files for the stimuli used in this thesis are available at http://people.pwf.cam.ac.uk/kjo21

ABSTRACT

Although familiarity with a talker or accent is known to facilitate perception, it is not clear what underlies this phenomenon. Previous research has focused primarily on whether listeners can learn to associate novel phonetic characteristics with low-level units such as features or phonemes. However, this neglects the potential role of phonetic information at many other levels of representation. To address this shortcoming, this thesis investigated perceptual learning of systematic phonetic detail relating to higher levels of linguistic structure, including prosodic, grammatical and morphological contexts. Furthermore, in contrast to many previous studies, this research used relatively natural stimuli and tasks, thus maximising its relevance to perceptual learning in ordinary listening situations.

Experiment 1 demonstrated that the stress context of an atypical vowel can influence perceptual learning. 'Global' subjects, who heard centralised /ɪ/ in both stressed and unstressed syllables during training, were more likely than Controls to accept realisations such as [ˈdəskəʊ] as a real word (*disco*). There was no such effect for 'Stress-conditioned' subjects, who heard centralised /ɪ/ only in unstressed syllables during training. However, different patterns of learning for Stress-conditioned subjects and Controls provide tentative evidence that hearing the atypical pronunciation in unstressed syllables also affected listeners' responses.

Experiments 2a and 2b showed that listeners adapted to duration differences that reflect prosodic structure, and this knowledge influenced their interpretation of syntactic structure. Experiment 3 demonstrated listeners' sensitivity to morphological context; listeners who heard the prefix *re-* pronounced as [rɪ] during training demonstrated perceptual learning that was at least partially specific to *re-* prefixes in an intelligibility-in-noise task.

This research shows that listeners can update their phonetic representations in response to incoming information and its relation to linguistic-structural context. In addition, certain patterns of systematic phonetic detail were more learnable than others. These findings are used to inform an account of how new information is integrated with prior experience in speech processing, within a framework that emphasises the importance of phonetic detail at multiple levels of representation.

CHAPTER 1
INTRODUCTION

1.1 OVERVIEW

In speech, there is not a one-to-one relationship between the acoustic signal and linguistic form. The precise articulation used by a speaker depends on a range of factors including their accent, rate and style of speech, and emotion, as well as the phonetic context of the sound and its position in the prosodic structure. The acoustic signal is also influenced by the physical characteristics of the speaker, such as their vocal tract shape, vocal fold properties and lung capacity, which may change with age and state of health. Background noise, echoes or electronic distortions can introduce further complexity to the speech signal. Yet, despite this variation, speech perception is remarkably robust: listeners are nearly always able to understand the meaning of their interlocutor.

The mechanisms that underlie this perceptual ability are still not well understood. It has proven difficult to identify acoustic invariants corresponding to abstract linguistic units such as features, phonemes or words, which are often assumed to be the basis of speech processing (e.g., McClelland & Elman, 1986; Norris, McQueen & Cutler, 2000; Stevens, 2002). In response to the 'problem of invariance', speech perception researchers have increasingly adopted a dynamic approach, moving away from attempts to identify invariant cues to phonological units, and towards models in which both the immediate context of the utterance and the past experience of the listener play a greater role (e.g., Goldinger, 1998; Hawkins & Smith, 2001; Johnson, 1997, 2006; Lachs et al., 2003; Pierrehumbert, 2003, 2006).

This more flexible approach to phonetic categories is supported by evidence that phonetic detail can be associated with multiple levels of representation, including higher levels of linguistic structure than the word and social dimensions of communication such as turn-taking, stylistic or talker-specific information. There is also considerable evidence to support the plasticity of phonetic representations; for example, familiarity with a particular talker, accent or unusual type of speech can facilitate perception.

However, little is known about the plasticity of perception in the context of multiple levels of representation. For example, to what extent can listeners adapt to a novel association

between phonetic cues and prosodic or syntactic structure? Is such adaptation possible through 'everyday' exposure to a talker (as opposed to lab-based training)? And are there constraints on the learnability of associations between certain phonetic cues and particular contexts or structures?

This thesis addresses these questions by investigating perceptual learning of context-sensitive phonetic detail. It uses relatively natural speech to explore whether listeners can learn to associate systematic phonetic detail with particular contexts or structures, and thus aims to elucidate the phonetic representations that underlie everyday speech perception.

1.2 CHAPTER OUTLINE

This chapter describes the background to the dynamic, context-sensitive model of speech perception that provides the theoretical framework for this research. Section 1.3 discusses the influence of multiple levels of representation on speech perception, while Section 1.4 reviews current knowledge of the levels of representation that can become associated with phonetic detail during perceptual learning. Finally, Section 1.5 introduces the research questions addressed in this thesis.

1.3 MULTIPLE LEVELS OF REPRESENTATION

Models of speech perception must be able to account for the abstract categories that enable listeners to understand speech despite the phonetic variation between speakers and styles of speech. However, after half a century of acoustic-phonetic research, invariant cues to linguistic units such as features, phonemes or syllables remain elusive. This suggests that the assumption of direct associations between linguistic units and the acoustic signal may be too simplistic. This is not a novel proposal: over sixty years ago, Firthian Prosodic Analysis (FPA) (Firth, 1948) emphasised the influence of structure and meaning on phonetic form. Unfortunately, this insight was largely ignored so far as its implications for models of speech perception were concerned, and only recently has the perceptual relevance of systematic phonetic detail relating to many levels of structure been given serious attention.

Sections 1.3.1–1.3.6 briefly review the evidence that phonetic detail reflects multiple levels of representation, and discuss the implications for models of speech perception.

1.3.1 Phonetic detail conditioned by local and long-domain phonetic context

A model which emphasises the significance of phonetic detail at multiple levels of representation need not reject the importance of phonetic detail relating to low-level phonetic categories such as features. There are plainly correlations between acoustic characteristics such as voice-onset-time (VOT) and the distinction between, for example, English *time* and *dime* which is usually termed a difference in the linguistic feature 'voice'. Similarly, the spectral centre of gravity is a reliable acoustic cue which relates to the difference between /s/ and /ʃ/, a difference in place of articulation which is often described in terms of the feature 'anterior' (e.g., Giegerich, 1992). The importance of acoustic cues to particular features is emphasised in approaches such as Stevens' Lexical Access From Features (LAFF) model (Stevens et al., 1992; Stevens, 2002), and (for a more limited set of features) by Zue's concept of 'robust features' (Zue, 1990).

Nevertheless, it is clearly inadequate to propose a model of speech perception which attempts to identify individual features (or phonemes as clusters of features) by their local acoustic characteristics alone, without reference to the immediate phonetic context. For example, the spectral centre of gravity that induces an /s/ or /ʃ/ percept varies depending on the surrounding vowel context, and this can cause a boundary shift such that lower frequencies are required to produce an /ʃ/ percept in the context of a rounded vowel than an unrounded vowel (Mann & Repp, 1980). Elman and McClelland (1988) demonstrated the pervasive influence of phonetic context on perception when they showed that even a lexically-restored segment can influence the identification of the following segment. Subjects were more likely to identify the first sound in an ambiguous [t/k]*apes* stimulus as /k/ when it was preceded by *Christma*[s/ʃ] than when it was preceded by *fooli*[s/ʃ]. Lexical information had influenced the perception of the ambiguous [s/ʃ] segment and triggered perceptual compensation for coarticulation.

Coarticulatory cues can also be perceptually useful over longer domains. Listeners can compensate for vowel-to-vowel coarticulation (e.g., Fowler, 1981; Alfonso & Baer, 1982), although such compensation may only be partial (Beddor et al., 2001). Long-domain resonance effects can also affect perception: identification of a missing /l/ or /r/ is better than chance when surrounding vowels and consonants in a carrier phrase are replaced by noise (West, 1999).

Therefore, although there are broad correlations between phonological features and acoustic characteristics, these correlations become more reliable when the local and long-domain

phonetic context is taken into account. It follows that phonetic cues are best interpreted with reference to local and long-domain phonetic context.

The following sections cite evidence that taking into account variation conditioned by broader linguistic-structural contexts (Sections 1.3.2–1.3.4) and wider communicative contexts (Section 1.3.5) can also be perceptually useful and affect the linguistic interpretation of an utterance.

1.3.2 Phonetic detail conditioned by syllabic context

The phonetic realisation of a segment is dependent in part on its position in the syllable, and this syllable-conditioned allophonic variation can influence perception at several levels of representation. Firstly, it is known to affect the assignment of word boundaries. For example, in British English, the sequence [greɪʔreɪn] is interpreted as *great rain* rather than the phonemically-identical sequence *grey train*, because [ʔ] is not a legal allophone of /t/ in the latter context. More subtle durational and spectral variation at word boundaries has also been shown to facilitate word segmentation in adults (Davis et al., 2002; Kirk, 2000; Quené, 1993; Shatzman & McQueen, 2006; Smith & Hawkins, 2000; Spinelli et al., 2007), and infants younger than 11-months old are sensitive to such variation (Jusczyk et al., 1999; Mattys & Jusczyk, 2001).

Secondly, listeners can use syllable-dependent phonetic variation to assist syllable segmentation within words. In a cross-modal priming study in Italian, Tabossi et al. (2000) found that hearing the onset of a word primed semantically-related visual targets only when it contained the appropriate syllable structure, so, for example, /sɒl/ from *sol.dato* (*soldier*) primed *guerra* (*war*), but /sɒl/ from *so.lare* (*solar*) did not. Tagliapietra et al. (2009) extended this finding to stressed syllables. At present, it is not clear whether subtle syllable-conditioned phonetic detail affects listeners' perception of words with contrasting syllable structure in English in the same way, because the role of the syllable is known to differ cross-linguistically. For example, Cutler et al. (1986) argue that English listeners rely less on within-word syllable segmentation than French listeners due to the high variability of syllable structure and the prevalence of ambisyllabicity in English.

Nevertheless, English listeners' perception of morphological boundaries can be influenced by phonetic detail that is associated with differences in syllable structure. Baker (2008) found that, in cases where differences in morphological structure drive differences in

syllabification, speech perception in noise is disrupted when phonemically-identical but morphologically-different material is cross-spliced (see Section 4.1.1 for further details).

Thus, phonetic variation conditioned by syllable structure can influence the perception of lexical, syllabic and morphological boundaries.

1.3.3 Phonetic detail conditioned by higher levels of prosodic structure

Higher levels of linguistic structure, such as intonational and phonological phrase boundaries, are also evident in the phonetic signal. Indeed, such structures are sometimes defined by prosodic characteristics such as phrase-final intonational contrasts (e.g., Pierrehumbert & Beckman, 1988). Prosodic position can affect many other aspects of articulation. For example, lengthening tends to occur at the end of a prosodic phrase (e.g., Cooper & Paccia-Cooper, 1980; Klatt, 1976; Wightman et al., 1992), the domain and degree of lengthening associated with primary phrasal stress may be modulated by prosodic boundaries (see Turk & White (1999) for more detailed discussion), while articulations are said to be 'stronger' at the beginning of prosodic domains, with greater domain-initial strengthening for structures higher in the prosodic hierarchy (Cho & Keating, 2001; Fougeron & Keating, 1997; Fougeron, 2001; White & Turk, 2010). The production of prosodic cues has been widely researched, and can be language- and accent-specific, and so a more extensive review is not attempted here.

Phonetic detail, both 'suprasegmental' and 'segmental', that is associated with higher levels of prosodic structure can affect the perception of syntactic structure and word boundaries. Listeners are known to use phrase-final lengthening as a cue to prosodic boundary position (e.g., Lehiste et al., 1976; Nagel et al., 1996; Scott, 1982; Speer et al., 1996), and pitch contours can also affect the assignment of phrase boundaries (Beach, 1991; Streeter, 1978). In addition, a recent study by Cho, McQueen and Cox (2007) showed that domain-initial strengthening in the form of higher peak amplitude and longer duration is also perceptually relevant.

1.3.4 Phonetic detail relating to grammatical function and meaning

A further level of representation that should be incorporated into models of perception is grammatical function. There are clear phonetic differences between the possible realisations of phonemically-identical words that differ in grammatical function. For instance, Lavoie (2002, p. 197) states that 'the preposition [*for*] has a range of very minimal realizations that are not attested or probably even acceptable realizations of the number [*four*]'. She

analysed both read and spontaneous American English speech, and found that the rime in *for* can be omitted (in an acceptable phonological context), whereas the number *four* is always realised with either an r-coloured vowel or an /r/. Similarly, the auxiliary *have* has many possible realisations, such as [hav həv əv v], but this range of reductions is not acceptable in minimally-different phonological forms that are content words, such as *ham, heave* or *Gav* (Foulkes, 2006).

Such phonetic differences between function and content words can affect perception: in a study of British English, Baker (2008) found that phonetic detail from an inappropriate grammatical context (function vs. content) could disrupt the perception of speech in noise (see Section 3.1.1 for further details).

There is also evidence that the precise meaning of a word can create more subtle phonetic differences. For example, /æ/ tends to be more raised in the word *hand* when it refers to a limb than in phrases such as *give a hand*, and /ɪ/ tends more centralised in the word *give* when it refers to abstract themes (e.g., *give a chance*) than when there is a meaning transfer (e.g., *give a pen*) (Hay & Bresnan, 2006).

The relationship between phonetic detail, grammar and meaning is discussed in more detail in Chapter 3.

1.3.5 Phonetic detail conditioned by interactional context

Phonetic cues can be associated not only with linguistic structure, but also with what might be termed 'social' aspects of communication – those aspects which relate more to the interactional context of a conversation. These are discussed in Sections 1.3.5.1–1.3.5.3.

1.3.5.1 Turn-taking cues

The production and perception of turn-taking cues is important for successful interactions. Prosodic cues are known to play a role in the perception of turn-endings (e.g., Couper-Kuhlen & Selting, 1996; Cutler & Pearson, 1986). In addition, research into the phonetics of 'talk-in-interaction' has identified systematic 'segmental' phonetic cues that are produced in particular interactional contexts. For example, Local (2003, p. 326) claims that, in Tyneside English, the aspiration of word-final voiceless plosives, in conjunction with centralisation of vowel qualities in the turn-final foot, is 'criterial for signalling turn transition'. Individual words can also be realised very differently depending on their interactional function. Local (2007) contrasts the phonetic realisation of 'holding *so*' in which a speaker continues to speak on the same topic, with 'trail-off *so*', which precedes a

change in talker. In contrast to 'holding *so*', 'trail-off *so*' tends to be quieter and lower in pitch than the accented part of the final foot of the preceding speech, it never has final glottal closure, and can be creaky throughout the whole syllable. The perceptual consequences of these production differences have not been explicitly investigated, but it seems probable that listeners take advantage of such cues to guide turn-taking in conversation.

This evidence suggests that discourse structure adds another necessary 'layer of representation' to a plausible model of speech perception.

1.3.5.2 Phonetic detail reflecting pragmatic and emotional meaning

Phonetic cues can be crucial in informing the listener about pragmatic differences in lexically-identical sequences. For example, a lowered pitch and relatively monotonic intonation might convey sarcasm in English (e.g., Cheang & Pell, 2008). More subtle affective aspects of meaning are also realised phonetically in the speech signal; for instance, high pitch might indicate fright (see Frick, 1985), while creaky phonation can be used to signal bored resignation in English (RP) (Laver, 1980). Nygaard and Lunders (2002) demonstrated that 'emotional tone of voice' can act to contextually constrain lexical access. They showed that listeners can make use of tone of voice to help disambiguate 'affective' vs. 'neutral' homophones such as *die* and *dye*, although they do not define the acoustic correlates of 'tone of voice'.

These affective cues are not restricted to pitch; segmental aspects of the speech signal can also be influenced by emotion. One such example is described by Laver (2003, p. 414):

> In English, and perhaps universally, incipient laughter is paralinguistically signalled partly by a medium-term articulatory setting which compresses the vowel space up and back towards the velum and the back wall of the pharynx, phonetically shifting vowels which are normally non-high and non-back slightly up and back in the vowel space. Hearing the first syllable or two of such speech leaves the listener uncertain about the linguistic phonemic values to allocate to the vowels heard, except by paralinguistically recognizing the incipience of the laughter by hearing the following syllables.

Fairly detailed pragmatic information can also be conveyed by the speech rate and degree of reduction employed by a speaker. Plug (2005) reports that discourse function influences the realisation of the Dutch word *eigenlijk* (which translates approximately as *actually* in

English). In a phonetic analysis of a corpus of conversational Dutch speech, he found that where *eigenlijk* is used to amend one's own prior talk, it tends to be fast and reduced; in contrast, when its function is to address problems arising from an interlocutor's talk, it is likely to be slow and unreduced. Plug also cites examples indicating that these phonetic characteristics extend beyond the word *eigenlijk* to phrases that perform the same pragmatic function.

Hawkins and Smith (2001) emphasise the importance of phonetic cues to pragmatic function. They describe how the linguistic meaning *I do not know* can be conveyed alongside detailed pragmatic information. If the fully expanded form is uttered, with pauses between the words (*I…do…not…know*), this is likely to be 'so rude that it can only be used when the listener does not seem willing to accept that the speaker really does not know' (p. 170). While, at the other extreme, in some circumstances, a highly reduced [ã̲ˀn:əᵘ] or [ə̰ə̰ə̰] (intonation not marked) can be enough to convey the meaning *I do not know*, probably alongside the implication that the speaker is not fully engaged with the topic, and is disinclined to help the listener find the answer. Between these extreme phonetic forms, a range of pronunciations are possible, conveying a greater or lesser degree of attention, exasperation and helpfulness on the part of the speaker. Importantly, Hawkins and Smith (p. 171) also highlight the importance of pragmatic context in correctly interpreting the acoustic signal:

> Because [listener] A sees that [speaker] B is deeply involved with her book, he will be less likely to interpret her [ə̰ə̰ə̰] as a sort of dysarthric grunt preparatory to a more helpful utterance. In other words, the meaning can be arrived at by linking the perceived multi-faceted (i.e. detailed) situation with perceived sound in a mutually reinforcing way…

The phonetic signal conveys pragmatic information, but, equally, it must be interpreted in relation to the wider pragmatic (real-world) context.

1.3.5.3 Phonetic detail reflecting talker-related contexts

A further 'social' aspect of communication that affects phonetic representations is the identity of the speaker or type of speaker. For example, expectations of vowel quality which are (presumably) derived from prior experience of social interaction can influence phonetic boundaries. Johnson, Strand and D'Imperio (1999) demonstrated that the gender

of a visually-presented face can shift the category boundary on an [ʊ]–[ʌ] auditory continuum spoken in a gender-ambiguous voice.

In addition, Hay et al. (2006b) found that the perceived age and social class of a speaker can affect subjects' perceptual decisions. They studied New Zealand English diphthongs /iə/ and /eə/ that are involved in a merger-in-progress. For subjects who maintain the distinction between the diphthongs in their own speech, judgements in two-alternative forced-choice (2AFC) task between words such as *air* and *ear* were more accurate when stimuli were accompanied by a photo of an older speaker, i.e. when listeners expected the difference between the diphthongs to be meaningful. In addition, the perceived social class of the speaker influenced accuracy for listeners with and without the merged variant.

Other studies have demonstrated that information concerning the regional background of a speaker can affect judgements of vowel quality, especially when they involve particularly salient or stereotyped differences that relate to that region (Hay et al., 2006a; Niedzielski, 1999). There is also considerable evidence that phonetic information associated with a particular talker can affect perception. This is discussed in more detail in Section 1.4.8.

1.3.6 Multiple levels of representation and models of speech perception

In summary, phonetic detail is systematically associated with many different contexts, both linguistic and 'social', and there is increasing evidence that listeners are able to use such information to facilitate perception. This demonstrates the inadequacy of models which advocate early abstraction away from the speech signal to relatively low-level phonological units such as features or phonemes, or even to words or probabilities of words (e.g., Motor Theory (Liberman & Mattingly, 1986); TRACE (McClelland & Elman, 1986); Shortlist (Norris, 1994); Merge (Norris et al., 2000); Shortlist B (Norris & McQueen, 2008); LAFF (Stevens, 2002)).

Firstly, such models do not take advantage of systematic phonetic information relating to other levels of representation and are therefore sub-optimal in their use of the information in the speech signal. For example, interpreting the speech signal with reference to contextual information may be important in making phonetic categories distinct enough to be easily discriminated. Pierrehumbert (1993, p. 264) neatly illustrates this for the distributions of /s/ and /z/ durations. Using data from a single speaker, she found that when prosodic position was not taken into account, the distribution of /s/ and /z/ durations

overlapped substantially. However, when split into word-initial, word-final and phrase-final tokens, /z/ was always shorter than /s/, thus making the phonetic categories more distinct.

Secondly, phonetic detail that relates to different levels of linguistic structure or types of social context will interact. That is, there are not independent phonetic cues that contribute separately to each distinction, but rather each context will affect the realisation of many phonetic cues (relating to duration, spectral characteristics, intonation and voice quality), and it is the combined effect of these influences (and all the other situational cues) that the listener must interpret. Therefore, a model which restricts itself to one (or just a few) levels of structure is not just incomplete, but may in some cases be misleading because such a model prevents reference to all the relevant contextual factors. A more integrated approach, in which phonetic detail is interpreted in the context of multiple levels of linguistic structure and interactional factors simultaneously, is necessary.

Such an approach is possible within exemplar-based models of speech perception, in which phonetic detail relating to every 'episode' is stored, along with many other relevant sensory aspects of that experience. In exemplar models, perception occurs when the current stimulus activates stored representations of similar past experiences. Therefore, within a broad conception of 'similarity', perception of a speech stimulus is achieved through activation of stored representations that are similar in terms of phonetic detail that reflects, for example, linguistic-structural context, talker, social context, and emotional context. In this way, the listener achieves a percept that includes the linguistic meaning of the utterance, alongside social and pragmatic implications.

The exemplar approach implies that perception is fundamentally dynamic – the perceptual system is constantly evolving on the basis of new experience. The speech signal is assumed to be processed in conjunction with other aspects of the listener's context. If any particular phonetic characteristic occurs at a level above chance in combination with any meaningful or relevant aspect of the listener's experience at the time, whether that is visual, linguistic or relating to other aspects of the interactional context, they are likely to become associated. In this way, phonetic representations are continually updated, and detailed phonetic variation can become associated with many levels of representation. Therefore, as well as accounting for the evidence discussed in Sections 1.3.1– 1.3.5 that listeners are sensitive to phonetic detail that relates to multiple levels of representation, exemplar models also predict a strong role for context-sensitive phonetic information in the acquisition and adaptation of phonetic categories.

Despite this prediction, there has been relatively little research into the adaptation of phonetic categories in the context of phonetic information that is associated with multiple levels of representation. Consequently, little is known about the extent to which listeners can learn to associate novel phonetic detail with linguistic contexts (such as those discussed in Sections 1.3.1 to 1.3.4) and non-linguistic contexts (such as those discussed in Section 1.3.5). This thesis begins to explore the role of context-sensitive phonetic information in the adaptation of phonetic categories by investigating perceptual learning of novel context-sensitive phonetic detail by adults.

Before outlining the research questions more precisely, Section 1.4 reviews what is known about the levels of representation which can become associated with phonetic detail during perceptual learning. In this thesis, perceptual learning is defined as 'experience-induced changes in the way perceivers extract information' (Kellman & Garrigan, 2009; but see also Samuel & Kraljic, 2009). This relatively broad view of 'perceptual learning' includes studies of second language acquisition, long-term adaptation, and language acquisition by children, as well as perceptual learning by adults under laboratory conditions.

1.4 WHAT CATEGORIES CAN BECOME ASSOCIATED WITH PHONETIC DETAIL DURING PERCEPTUAL LEARNING?

1.4.1 Evidence for perceptual learning at multiple levels of representation

This sub-section addresses the question of what categories can become associated with phonetic detail by discussing evidence that familiarity with a particular type of speech can involve learning associations with various levels of structure.

Familiarity with an accent can include knowledge of lexical and sublexical phonetic cues. It has been shown that the processing of isolated words is disrupted by an unfamiliar accent (Adank & McQueen, 2007). In addition, most listeners are consciously aware of certain segmental differences associated with other accents – for example, that [p^hɪn] is a realisation of *pin* in SSBE, but of *pen* in New Zealand English.

Studies of sine-wave speech confirm that there is a segmental or sub-segmental element to familiarity with a talker. Remez et al. (1997) showed that listeners who were very familiar with particular talkers through everyday interactions could identify these talkers from samples of sine-wave speech. Sine-wave speech preserves the time-varying formant structure of speech while eliminating indexical information such as fundamental frequency, harmonic structure and characteristic intonation patterns. Therefore, Remez et al. argue that

listeners can recognise talkers from fine-grained individual differences in phonetic realisations, such as allophonic characteristics, talker-specific differences in coarticulation, or characteristic dialectal properties of segments.

Liss et al. (2002) investigated what can be learned during adaptation to a talker by investigating adaptation to two types of dysarthric speech: ataxic dysarthria, which is characterised by a perceptually rapid rate of speech and monotonicity, and hypokinetic dysarthria, which is characterised by a slow rate and an even syllabic stress. For both types of dysarthric speech, listeners who heard the same type of dysarthria in training and testing achieved the highest intelligibility score in a sentence transcription task, followed by listeners who had heard the other type of dysarthria, followed by the no-familiarisation condition. No difference was found in the pattern of word boundary errors between the groups, but there was some evidence to suggest between-group differences in phonemic errors. The authors suggest that 'the cognitive-perceptual source of benefit may not lie in the mapping of suprasegmental patterning' (p. 3028). However, there are many other indexes of suprasegmental learning than word boundary errors, so this is by no means conclusive.

Certainly, other evidence indicates that listeners can adapt to phonetic detail related to levels of structure above the word. Nygaard and Pisoni (1998) familiarised listeners with 10 voices reading isolated words, and then gave them an intelligibility test in which they had to identify isolated words presented in noise. Results showed a talker-familiarity effect: listeners could better identify novel words masked by noise when they were spoken by talkers whom they had heard in an earlier training session. Furthermore, when listeners were trained with sentence-length utterances, the intelligibility of words presented in sentences spoken by familiar talkers improved. Crucially, however, learning did not generalise from training with sentences to the identification of isolated words. This indicates that the training material – sentences or isolated words – influenced which talker-specific cues were learned, and therefore listeners must be sensitive to suprasegmental differences between talkers. This conclusion is further supported by the results of Weil (2001): training American listeners with one Marathi-accented talker using a variety of tasks (word identification, sentence transcription, and comprehension of short prose passages) generalised to facilitate comprehension of a novel Marathi-accented talker in a sentence transcription task but not a word transcription task.

Adaptation to time-compressed speech suggests that learning can relate to a relatively abstract mapping between the signal and rhythmic units, rather than highly specific acoustic-phonetic properties. Training with time-compressed speech can generalise across talkers and rates of compression, and also across languages which share certain rhythmic properties (Dupoux & Green, 1997; Sebastián-Gallés et al., 2000).

Therefore, phonetic detail can become associated with both segmental and suprasegmental representations during perceptual learning, and what is learned can vary considerably depending on the task and the phonetic cues available in the signal. The following sections (1.4.2–1.4.8) review research that can shed light on the specific phonetic cues that listeners can learn to associate with particular contexts.

1.4.2 Perceptual learning at sub-lexical and sub-segmental levels of structure

Many studies demonstrate that at least some of what is learned from exposure to a talker is at the level of sublexical representations. For example, Norris et al. (2003) demonstrated that exposure to an ambiguous segment in a lexical decision task could influence phoneme categorisation. In a between-groups study, one group of subjects completed a lexical decision task which included phonetically unambiguous /s/-final words and also /f/-final words in which the /f/ was replaced by an ambiguous sound between [f] and [s]. Another group of listeners heard phonetically unambiguous /f/-final words, while the final fricative of /s/-final words was replaced by the ambiguous [f/s] fricative. Following training, there was a boundary shift in listeners' identification of sounds on an [ɛf]–[ɛs] continuum: listeners who had been exposed to the ambiguous fricative in /f/-final words were more likely to categorise ambiguous sounds as /f/ than those who had heard it in /s/-final words. Thus, lexical knowledge influenced the interpretation of a sound, leading to a shift in the boundary between segments which was measurable in responses to a non-word–non-word continuum, clearly demonstrating a sublexical locus of perceptual learning. In support of this conclusion, McQueen et al. (2006) showed that the same training procedure also influenced the interpretation of novel words not used in training, as assessed by a cross-modal priming task.

Learning at a featural level of representation has also been demonstrated. Using the same paradigm as Norris et al. (2003), Kraljic and Samuel (2006) investigated generalisation of perceptual learning across place of articulation. Following a lexical decision task which included sounds (mixed from natural tokens) that were ambiguous between /d/ and /t/, listeners categorised sounds on [ada]–[ata], [ɪdɪ]–[ɪtɪ] and [aba]–[apa] and [ɪbɪ]–[ɪpɪ]

continua. Results showed a boundary shift across both places of articulation, indicating that learning was not specific to the segments heard during training, and may have related to the characteristic way the talker realises the 'voice' distinction (linguistically, a distinctive feature).

This is consistent with evidence from children having therapy for phonological disorders; training a child to produce a voicing contrast at the bilabial place of articulation in nonsense syllables has been shown to generalise to alveolar and velar places of articulation (McReynolds & Bennett, 1972). Further evidence for a featural level of learning comes from therapy with children who substituted stops instead of most continuants (e.g., [t] for /s/): after being successfully trained to produce [s] and [θ], they were able to produce [ð], [z] and [ʃ] (Costello & Onstine, 1976).

Maye et al. (2008) also demonstrate sub-lexical perceptual learning, although the precise locus of learning is uncertain. Listeners heard a story in which the (synthetic) speaker had lowered front vowels. After hearing the story, listeners became more likely to accept lowered front vowel pronunciations, such as *wetch* for *witch*, as real words in a lexical decision task. A control experiment indicated that listeners had not simply broadened their front vowel categories – items with raised front vowels were not more likely to be accepted as words. However, it is not clear from this study whether 'lowering' was learned at a segmental level (i.e. about each vowel individually), or whether listeners learned that front vowels as a class were lowered.

As well as adjusting a phonetic representation by adapting the 'value' of a particular characteristic (e.g., VOT or formant frequency), listeners can also adapt phonetic representations by changing the 'weight' given to different acoustic characteristics. Francis et al. (2000) successfully trained listeners to adjust the cue to which they pay most attention when categorising place of articulation in stops. They created synthetic stimuli with cooperating and conflicting cues by manipulating the spectrum of the burst release and the origin of the formant transitions. One group of listeners was trained to classify the stops according to the burst release spectrum, and the other according to the formant transitions. Results showed that both groups shifted towards using the specific property they had been trained on.

These studies demonstrate the plasticity of representations at the sublexical level, with particular phonetic cues becoming associated with segmental and featural contrasts, and possibly other sublexical units. However, they do not consider whether listeners can learn

about the broader linguistic or social contexts in which a particular cue occurs. Sections 1.4.3–1.4.10 discuss perceptual learning of phonetic detail that is conditioned by such contexts.

1.4.3 Phonetic detail conditioned by adjacent phonetic context

Infants can rapidly learn that phonetic characteristics are dependent on the adjacent phonetic context. In a between-subjects design, White et al. (2008) familiarised infants with an artificial language in order to investigate whether they could learn that stops (for Stop subjects) or fricatives (for Fricative subjects) are voiceless following a voiceless consonant but voiced in other contexts. In a 'determiner plus noun' sequence exemplifying this allophonic constraint, infants were exposed to *na* or *rot* 'determiner' monosyllables before disyllabic words (e.g., *na bevi* and *rot pevi* (Stop subjects); *na zuma* and *rot suma* (Fricative subjects)). Stop subjects also heard disyllabic words beginning with voiced and voiceless fricatives after both *na* and *rot* (e.g., *na zuma*, *na suma*, *rot zobi*, *rot sobi*); and Fricative subjects heard voiced and voiceless stops after both *na* and *rot* (e.g., *na bevi*, *na pevi*, *rot bogu*, *rot pogu*). In the test task, all infants heard the same 'determiners' paired with new 'nouns'. Half of the nouns began with stops and half with fricatives, and all occurred voiced after *na* and voiceless after *rot* (e.g., Stop-initial sequence: *rot poli*, *na boli…* and Fricative-initial sequence: *rot sadu*, *na zadu…*). Both 8.5-month-olds and 12-month-olds showed a preference for sequences that, on the basis of the allophonically-conditioned variation heard during familiarisation, exemplified one novel alternating noun, i.e. Stop subjects preferred stop-initial disyllables, while Fricative subjects preferred fricative-initial disyllables.

However, the authors note that this result could be due to the infants learning the transitional probability between the determiner and the following phoneme, rather than to infants learning to functionally group voiced and voiceless segments in that particular context. Follow-up experiments removed transitional probability cues from the test task by exposing infants to novel nouns without determiners (e.g., *poli*, *boli…*). In this situation, 8.5-month-olds failed to show a preference, while 12-month-olds still preferred the novel alternating noun. On this basis, White et al. suggest (p. 255) that infants first become sensitive to the probabilistic relationship between sounds and their conditioning contexts, and later develop the ability to group similar sounds which occur in complementary conditioning contexts.

White et al.'s study demonstrates perceptual learning of an allophonic rule from brief exposure (2 minutes). However, the different behaviour of infants at different stages of

development suggests that prior knowledge and/or developmental progress can affect what is learned from a stimulus. This highlights the caution necessary in generalising conclusions from infant to adult perceptual learning.

A similar artificial language paradigm was used to test the learning of allophonic rules conditioned by adjacent voicing context in adults (Peperkamp & Dupoux, 2007). In a between-subjects phrase–picture matching task, subjects were familiarised with 'determiner plus noun' sequences in which the voicing of either initial stops or initial fricatives was dependent on the preceding voicing context, while there was no such dependency for the other obstruents (either stops or fricatives). In a subsequent word–picture matching task, subjects were more likely to ignore a voicing difference for segments heard in the allophonic condition (i.e. dependent on the voicing context) than in the phonemic condition (i.e. independent of the voicing context), for both familiar items (heard during training) and novel items.

Interestingly, the learning did not generalise to novel place of articulation, indicating that perhaps in this instance subjects were learning on a segment-by-segment basis, rather than about a natural phonological class (either stops or fricatives). This is supported by a subsequent experiment showing that listeners were able to learn about voicing differences in unnatural allophonic groupings. That is, rather than the allophonic rule applying to either the class of stops or the class of fricatives as a whole, some stops and fricatives were subject to the constraint while others were not. This contrasts with the 'feature-based' learning demonstrated by Kraljic et al. (2006, see Section 1.4.2), and therefore indicates that there may not be a 'default' locus of perceptual learning; instead, different stimuli or tasks may produce perceptual learning at different levels of representation.

In another study demonstrating adaptation to phonetic characteristics conditioned by adjacent context in adult listeners, Dahan et al. (2008) used eye-tracking to monitor sensitivity to phonetic context in adapting to an existing (non-artificial) accent. They investigated listeners' response to the American dialect in which the vowel /æ/ raises to a vowel approaching [ɛ] before /g/, but not before /k/. Subjects were not given a separate training session with the accent, but rather just experienced the isolated '*bag*-like' or '*back*-like' words in the eye-tracking task. Results showed that subjects were able to make use of the accent to reduce lexical competition between words like *bag* and *back*. However, subjects' prior familiarity with the talker's dialect was not assessed, so it is unclear to what

extent they were learning a novel cue, and to what extent they were drawing on prior experience of the accent.

Evidence of learning of phonetic detail conditioned by surrounding phonetic context is not always found. Kraljic et al. (2008a) investigated adaptation to an ambiguous sound between /s/ and /ʃ/ when context-independent (all instances of /s/ were replaced by the ambiguous sound), or context-dependent (only /s/ in *str* was replaced; /s/ in other contexts was not replaced by the ambiguous segment). Subjects were familiarised with one accent or the other using a lexical decision task. A subsequent boundary shift occurred for subjects trained on the context-independent sound on both [asi]–[aʃi] and [astri]–[aʃtri] continua. However, subjects trained on the context-dependent sound showed no boundary shift, even on the [astri]–[aʃtri] continuum, which replicated the *str* context in which the unusual variant had been heard during training.

This confirms that the context in which an ambiguous segment occurs significantly affects what is learned, but differs from the studies cited above in that in this case the context-sensitive variant appeared not to trigger adaptation. The authors offer three possible explanations for this.

Firstly, the occurrence of the ambiguous /s/–/ʃ/ variant in the context of *str* does occur dialectally (e.g., in some speakers of New York English), so the authors suggest that prior familiarity with the dialect may have limited the degree of learning. However, this dialectal variation could have led to the opposite prediction, that adaptation should be particularly rapid. The results of Dahan et al. (2008), described above, suggest that prior knowledge of a dialect may not inhibit adaptation.

Secondly, the context-conditioned variation could be attributed to coarticulation, and this could potentially limit its influence on perceptual learning.

Thirdly, it may reflect phonotactic knowledge. /ʃtr/ is not a legal onset in English, and is never contrastive with /str/, so it is possible that no learning takes place because the 'ambiguous sound' would never lead to word identification difficulty. This could relate to the depth of processing – less ambiguity in the lexical decision task may have led to shallower processing and hence less adaptation.

An alternative possibility is that the phonotactic constraint on /ʃtr/ may affect the nature of the categorisation task. In the [astri]–[aʃtri] continuum, the presence of [ʃ] may encourage syllabification as aʃ.tri (i.e. as two phonotactically legal syllables), thereby making the

phonological context of [ʃ] in testing different from during training. If this is the case, then the test task is an inappropriate assessment of whether context-sensitive perceptual learning occurred. Whether this explanation is valid may depend on the duration of the ambiguous segment – a relatively short segment would be more likely to be perceived as syllable-final.

It is also possible that context-sensitive constraints may take longer to learn than context-independent ones. Longer adaptation periods for 'second-order constraints' were noted by Dell et al. (2000; see Section 1.4.4).

In summary, there is some evidence of perceptual learning of phonetic detail conditioned by the adjacent phonological (or phonetic) context, and that what is learned depends to an extent on prior knowledge. However, relatively brief exposure to a variant in a restricted phonological context does not seem necessarily to induce context-sensitive learning, and may be affected by complex factors relating to the specific phonetic characteristics and task involved.

1.4.4 Phonetic detail conditioned by syllabic structure

There is some evidence to suggest perceptual learning of phonetic detail conditioned by syllabic structure. Indeed, learning to relate phonetic detail to syllabic structure may be more natural in some training situations than learning to associate phonetic detail with phonemes (in the sense of properties that generalise across positions within the syllable). For example, in a study of second language learning, Morosan and Jamieson (1989) trained monolingual adult Canadian francophones to distinguish /θ/ from /ð/ using synthetic /Cʌ/ stimuli. At test, learning generalised to natural tokens of /Cʌ/, but failed to generalise to natural tokens with the target segment in word-medial and word-final positions in both real words and nonsense words. Therefore, perceptual learning was not at a phonemic level, and appeared to relate to the segment's position within the syllable. This conclusion must be somewhat tentative, as it is also possible that listeners had learned the contrast only in the context of vowel /ʌ/, or simply that the learning was most evident in the test tokens that were acoustically most similar to those heard during training. Moreover, /Cʌ/ is not an acceptable phonological word in English – a phonologically long vowel or short vowel and coda consonant(s) are necessary (except in function words e.g. *the*) – and this could potentially have influenced ease of learning and/or generalisation to real word contexts. Nevertheless, these results highlight problems with treating the phoneme as the default domain of perceptual learning, and suggest instead that perceptual learning is sensitive to syllabic structure.

Several production studies have demonstrated the learnability of constraints relating to syllabic structure. Dell et al. (2000) used speech errors to investigate the learning of experiment-wide constraints in which certain segments were permitted only in either onsets or codas. Subjects were asked to recite four-syllable sequences (e.g., *feng keg hem nes*) in time with a metronome. Results showed that listeners' speech errors followed the 'pattern' they had been exposed to in the word list, i.e. listeners who had only been required to produce /f/ in onset position throughout the experiment would tend to produce /f/ as a speech error more frequently in onset position than coda position.

A subsequent experiment by Dell et al. showed that listeners were also able to associate consonants with particular positions in the syllable in the context of a particular vowel. For example, where /f/ occurred in onset position only before /æ/ and in coda position only after /ɪ/, speech errors reflected this pattern, although these 'second-order' constraints took longer to learn.

Onishi et al. (2002) used a production task to test for contingencies learned through auditory experience (as opposed to production experience as used in the Dell et al. (2000)). Listeners heard CVC syllables which had constraints on consonant position, and this was followed by a speeded repetition task containing 'legal' syllables (which followed the constraints) and 'illegal' syllables (which did not). Results showed that legal syllables were repeated faster than illegal syllables. A pattern in which the legal consonant position depended on the adjacent vowel was also learnable, but there was no evidence of learning of a contingency in which the legal consonant position depended on the talker. This last result suggests that perhaps certain types of constraint are difficult to associate with a specific talker, at least in certain learning environments. It may reflect the associations that prior experience suggests are plausible; a contingency in which certain talkers do not produce certain phonemes in a particular position within the syllabic structure may be unlikely, especially for a native speaker. However, in a situation in which the listener knew they were listening to a foreign accent, perhaps such a contingency would be learnable. The role of prior experience in perceptual learning is discussed further in Chapter 5.

Finally, it has been shown that infants are sensitive to voicing conditioned by syllabic structure. Saffran and Thiessen (2003) exposed one group of 9-month-olds to synthetic nonsense words in which syllable-initial consonants were voiceless and syllable-final consonants were voiced (e.g., *todkad*), while another group heard the phoneme order reversed to create the opposite pattern of voicing (e.g., *dakdot*). In a subsequent task, using

four new words (two from each voicing pattern), infants listened longer to words which failed to conform to the pattern heard during training. A follow-up experiment tested whether the babies were learning a pattern relating to phonemes or features by investigating the learning of stimuli with consistent restrictions on phonemes but not voicing i.e. can infants learn that syllables can begin with /p/, /d/ and /k/ and end with /b/, /t/ and /g/ (or the reverse)? In this experiment, the infants showed no preference for the pattern heard during training. This suggests that hearing a consistent pattern in terms of features (and therefore also in acoustic terms) can be important for learning, and so associating phonetic cues with phonemes may not be the default position.

However, this contrasts with Peperkamp and Dupoux (2007; see Section 1.4.3) who found that listeners did not generalise across place of articulation, and therefore seemed to relate the voicing contingency to individual segments rather than learning a distinctive feature that applied to the phonological class of stops. This could be due to differences in the training stimuli or the conditioning context affecting what was learned. Alternatively, adults may have stronger representations for individual segments than infants (potentially influenced by literacy in an alphabetic system), and this prior knowledge may make them more likely to learn segment-based associations. For infants, the acoustically-consistent natural class may have been a relatively more prominent category.

In summary, there is some evidence for context-sensitive perceptual learning relating to syllabic structure. However, the most convincing examples come from infant studies and production studies, and should therefore be interpreted cautiously in terms of what they can tell us about phonetic representations in speech perception in adults. In addition, there are further indications that prior knowledge affects which contexts become associated with a phonetic cue (see Chapter 5 for a more detailed discussion).

1.4.5 Phonetic detail associated with phrase boundaries

At a higher level of linguistic structure, there is evidence of perceptual learning of phonetic detail dependent on the presence or absence of a phrase boundary. Scott and Cutler (1984) examined alveolar flapping, which is inhibited over phrase boundaries in American English (Cooper & Paccia-Cooper, 1980). For example, in the sentence *For those of you who'd like to eat, early lunch will be served*, alveolar flapping is not permitted in *eat*. In contrast, in *For those of you who'd like to eat early, lunch will be served*, alveolar flapping is optional. Alveolar flapping is not a feature of most British English accents and was rarer at the time of this experiment than it is now, and so long-term adaptation to a phrase boundary cue was

investigated by examining responses of long-term British residents in the US. Results showed that while speakers of American English used alveolar flapping to aid comprehension of ambiguous sentences, it was not used by speakers of British English in the UK. However, British English speakers who were long-term residents in the US used the cue to some extent.

Although this demonstrates a degree of perceptual learning of a phrase boundary cue, it also highlights some of the limitations of perceptual learning; even after years of living in the US, the performance of British English speakers did not equal that of native American English speakers.

1.4.6 Phonetic detail associated with pragmatic context

As discussed in Section 1.3.5, contexts other than linguistic structure can influence the phonetic detail of the speech signal. Kraljic et al. (2008b) demonstrated the importance of pragmatic context in what is learned from training. They assessed perceptual learning in an audiovisual condition in which an ambiguous [s/ʃ] replaced either /s/ or /ʃ/, in different conditions. This ambiguous pronunciation was contingent on the speaker having a pen in their mouth, while non-ambiguous /s/ and /ʃ/ occurred when the speaker did not have a pen in their mouth. A control condition included ambiguous and non-ambiguous instances of /s/ (or /ʃ/) that were not contingent on the pragmatic context (i.e. they were random). Results of categorisation task on a /s/–/ʃ/ continuum (presented only the auditory modality) showed a boundary shift in the control conditions but not in the conditions where the atypical variant was dependent on the presence of a pen in the speaker's mouth.

Kraljic et al. (p. 336) conclude that:

> The [perceptual] system integrates available cues about whether a variation is characteristic of the speaker who is producing it or an incidental consequence of some other factor. If the variation seems characteristic, the appropriate phonemic representation is restructured to accommodate it; if the variation seems incidental, no such restructuring occurs. The process of rapidly recognizing and extracting invariance is guided by pragmatic attributions about the variation's source.

The conclusion that different expectations about the source of variation can affect processing is consistent with Onishi et al. (2002) (see Section 1.4.4), and is an important demonstration that the extent of perceptual learning can be affected by pragmatic context.

However, the authors' claim that 'no restructuring occurs' when the variation is 'incidental' is less obviously correct. This experiment assessed perceptual learning using an auditory-only continuum, and did not include an audiovisual assessment task to investigate whether listeners had learned to associate the ambiguous [s/ʃ] with the pragmatic 'pen-in-mouth' context. An audiovisual task that assesses listeners' categorisation on a /s/–/ʃ/ continuum when presented with the talker with a pen in their mouth would have been necessary in order to determine whether listeners in the 'pen-in-mouth' condition did in fact restructure their phonetic representations, i.e. whether these listeners learned how the speaker says /ʃ/ (or /s/) with a pen in their mouth.

Consequently, this experiment provides evidence that listeners are sensitive to pragmatic context during perceptual learning, but it remains unclear whether listeners can learn to associate phonetic detail systematically with a particular pragmatic context.

1.4.7 Cross-modal variation: Phonetic detail associated with visual context

Perceptual learning experiments confirm the importance of integrating auditory and visual information in phonetic representations. In an experiment based on the McGurk effect (McGurk & MacDonald, 1976), Bertelson et al. (2003) exposed listeners to ambiguous tokens of a synthetic sound intermediate between /ada/ and /aba/ dubbed onto videos of the speaker articulating either /ada/ or /aba/. After exposure to the /ada/ visual stimuli, there were more /ada/ responses in a subsequent auditory forced-choice identification task; conversely, there were more /aba/ responses after exposure to the /aba/ visual stimuli. This suggests that, through the visual cue, the ambiguous sound had become associated with either an /ada/ or an /aba/ interpretation. Thus, 'phonetic' categories can be updated with reference to non-auditory dimensions.

This conclusion is further supported by evidence from second language acquisition. Japanese listeners learn a labial–labiodental contrast (between /b/ or /p/ and /v/) more effectively with audiovisual training than purely auditory training (Hazan et al., 2005), indicating that visual information may contribute to the development of accurate category representations in adults. However, audiovisual training was no more effective than auditory training for the /r/–/l/ contrast. This suggests that the degree of cross-modal integration during category-learning is dependent on the informativeness of the visual cues – for native English speakers categorising the stimuli using only visual cues, accuracy was 92% for /v/ (vs. /p/ or /b/), but only 71% for the /l/ vs. /r/ contrast. Therefore, where visual information is especially informative about the phonetic category (e.g., in distinguishing a

labial from a labiodental segment), it is more likely to become a significant 'dimension' of the newly-acquired phonetic representation.

These studies demonstrate that perceptual learning takes place with reference to the visual context. A more detailed discussion of audiovisually-induced perceptual learning can be found in Samuel and Kraljic (2009).

1.4.8 'Talker-specific' perceptual learning

The talker can be viewed as another 'non-linguistic context' that can constrain phonetic characteristics. Detailed encoding of the speech signal is demonstrated by perceptual learning studies that show the retention of phonetic variation that characterises a talker. Several studies have demonstrated a processing advantage for words heard in a familiar voice; Goldinger (1996) and Nygaard and Pisoni (1998) found that subjects were better able to identify words masked by noise when presented in the same voice as in an earlier training session. Word recognition in noise is also better when the talker remains the same across a block of trials than when multiple talkers are heard (Mullennix et al., 1989; Sommers et al., 1994), and there is increased priming in an auditory stem-completion task when the stem is presented in a familiar voice (Schacter & Church, 1992).

Recent evidence from voice identification in forensic phonetics provides further support for the encoding of subtle talker-specific phonetic detail. For example, listeners are reasonably accurate at voice identification when identifying a target voice at telephone quality after prior exposure at studio quality; however, their performance is significantly poorer when identifying voices from studio-quality recordings after previously hearing the talker only in telephone-quality recordings (Hudson, McDougall & Nolan, 2010). This suggests that listeners can 'filter' the studio recordings to identify the telephone voices, but are unable to reconstruct talker-specific phonetic detail from telephone-quality exposure, and thus perform less well at voice identification.

However, although these studies reveal learning of talker-related phonetic detail, it is not clear precisely what phonetic characteristics are encoded and used in perception. The following studies aimed to identify the locus of perceptual learning more specifically.

Allen and Miller (2004) investigated whether listeners are sensitive to individual differences in VOT. Using synthetic stimuli, they demonstrated that listeners can learn to associate a relatively short or relatively long VOT with a particular talker. Listeners were able to identify which of two variants of *town* spoken by a particular talker was consistent

with their experience of that talker during training. Furthermore, this knowledge of talker-specific VOT generalised to the novel word *time*. Therefore, perceptual learning can be highly context-specific, relating to how a particular talker realises a particular phonetic property.

Listeners also seem able to learn about talker differences in the realisation of word boundaries (Smith, 2007, submitted). The intelligibility-in-noise of sequences with ambiguous word boundaries (e.g., *Pat's awed* or *Pat sawed*) was tested before and after familiarisation with a single speaker saying sentences of this type. Results showed that subjects who heard the same voice in testing as in familiarisation showed more improvement than those who heard a different voice, indicating perceptual learning of the way a specific talker signals certain word boundaries.

However, as the following example illustrates, it is not always easy to verify that the 'context' associated with phonetic detail is a particular talker, rather than a more general acoustic-phonetic similarity of the test stimuli and the training stimuli, because talker and acoustic similarity are confounded.

Using the paradigm established by Norris et al. (2003), Eisner and McQueen (2005) describe a shift in a phoneme boundary that they claim to be talker-specific. Following an auditory lexical decision task including an ambiguous [f/s] fricative, there was a boundary shift in listeners' identification of sounds on an [ɛf]–[ɛs] continuum based on the same talker. This boundary shift did not occur when the [ɛf]–[ɛs] continuum was from a novel talker, supporting their claim of talker-specific perceptual learning. However, there was also a condition in which the test continuum was created from the vowel of a novel talker spliced onto fricatives from the talker heard during training. A boundary shift was present in this condition. This result is informative because listeners perceived these stimuli as being produced by a novel talker. This implies that the lack of a perceptual learning effect when the continuum was wholly from a novel talker was unlikely to be due to high-level influences relating to the recognition that a talker is novel. Rather, the boundary shift may be dependent on the acoustic-phonetic similarity of the training and test stimuli at the perceptually-relevant level (which, in a phoneme categorisation task, will be segmental – in this case, the fricative).

Stimuli produced by the same talker are likely to be more acoustically similar than those which are not, making it difficult to separate learning effects that relate to acoustic similarity from effects that are associated specifically with the identity of the talker.

Nevertheless, there is evidence that acoustic-phonetic similarity may be a significant factor in perceptual learning about 'talkers'.

Kraljic and Samuel (2005) investigated whether a shift in the /s/–/ʃ/ boundary generalised across talkers. An auditory lexical decision task was used to familiarise listeners with a talker, and a phoneme categorisation task assessed learning. An interesting asymmetry was found: training with a male talker did not induce a boundary shift on a /s/–/ʃ/ continuum from a female talker, while training with a female talker did generalise to testing on a continuum from a male talker. They suggest that this difference in talker-specificity can be explained by a closer inspection of the spectral characteristics of the fricatives in the training and test stimuli. The difference in spectral mean was greater between male training fricatives and female test items than between female training fricatives and male test items, hence the differences in patterns of generalisation may be explained by acoustic-phonetic similarity. In addition, Kraljic and Samuel (2006) found that a shift in the boundary between voiced and voiceless stops generalised across male and female talkers, and suggest that this is because the primarily temporal cues that distinguish these stops were shared by the talkers.

Further support for acoustic-phonetic similarity as a key component in 'talker-specific' effects comes from evidence that the degree of similarity between training and test voices affects the identification of words in noise and memory for previously heard words (Goldinger, 1996). This implies that, in at least some instances of 'talker-specific' perceptual learning, the phonetic characteristic is not specifically associated with the talker, but is associated, in a gradient fashion, with stimuli that sound like the talker. (In a more standard conversational situation, where the listener can see the speaker, the phonetic characteristics may also become associated with visual cues to talker identity, and thus, in practical terms, the learning is more talker-specific.)

In conclusion, fine-grained phonetic information that reflects variation between talkers is learned by the listener from relatively brief exposure, and can affect the perception of stimuli that share these phonetic characteristics.

1.4.9 Perceptual learning of sociophonetic variation

Recent work has investigated perceptual learning of sociophonetic variation. Docherty et al. (2008) demonstrated that adult listeners can learn to associate an intervocalic glottal variant of /t/ with a group of talkers. Listeners were trained by exposure to isolated word stimuli

that systematically associated intervocalic [t] with 'Tribe 1' and intervocalic [ʔ] with 'Tribe 2'. The test task was to classify words as being produced by a speaker of Tribe 1 or Tribe 2. Listeners were able to acquire the association both when the glottal variant was categorically associated with a particular 'tribe' during training, and when the association was probabilistic (there was a strong tendency for it to be produced by one tribe and not the other).

However, in a similar experiment, listeners were unable to learn a 'sociophonetic' association involving a diphthongised variant of /i/. So, although listeners can learn to associate a particular phonetic variant with a particular social group, there may be subtle influences governing how easy it is to learn about a particular variant. (This is discussed in more detail in Chapter 5.)

1.4.10 Context-sensitive representations: Summary of evidence from perceptual learning

Sections 1.4.1–1.4.9 have identified evidence that listeners can learn to associate systematic phonetic detail with different levels of linguistic structure – both segmental and prosodic – and also with non-linguistic contexts. However, that evidence is limited in several ways.

Firstly, despite evidence that perceptual learning can relate to contexts above the level of the segment (see Section 1.4.1), there has been little investigation of what lies behind such familiarity effects. Tightly-controlled studies investigating the learning of a particular phonetic characteristic have tended to focus on learning at relatively low levels of linguistic structure (Sections 1.4.2–1.4.4).

Secondly, a reasonable proportion of the evidence of sensitivity to contextually-conditioned phonetic detail comes from studies of infant learning and studies of production, and therefore the implications for speech perception in adults are not clear. In addition, for a number of studies, it is not clear exactly what has been learned by the listener, because full and systematic testing of the variant in many different contexts has not been carried out.

Thirdly, the one study of perceptual learning relating to phrase boundaries comes from a study of long-term adaptation to accent, and may rely on different mechanisms to the shorter-term perceptual learning effects discussed elsewhere. This does not make the phenomenon less important, but it does mean that it remains uncertain whether such effects can be learned within a short period of time, and that has implications for the degree of plasticity that should be incorporated into a model of speech perception.

Finally, context-dependent perceptual learning effects were not always found, and the reasons for this have yet to be established (Kraljic et al., 2008a; Onishi et al., 2002).

Clearly, there is much still to learn about the contexts that can become associated with novel phonetic detail, and the circumstances under which such learning occurs.

1.5 THEORETICAL MOTIVATION

This thesis aims to explore systematic phonetic detail relating to multiple levels of representation from a dynamic perspective. It investigates 'what is learnable' in terms of associating phonetic detail with particular contexts. It has two interrelated goals: to add to our knowledge of the levels of representation with which novel phonetic detail can become associated, and which must therefore be incorporated into a dynamic model of speech perception; and to add detail to the as yet rather vague notions of plasticity and stability that are (or need to be) fundamental to exemplar approaches. These goals are discussed in more detail below.

1.5.1 Elucidating multiple levels of representation in an exemplar approach

Section 1.3 presented evidence that phonetic detail that relates to multiple levels of representation can facilitate speech perception. Such evidence is consistent with exemplar models of speech perception, in which phonetic representations are based on episodic representations encoded in memory, and through this encoding, phonetic detail can become associated with many contexts (both linguistic and non-linguistic), as discussed in Section 1.3.6. However, different exemplar-based approaches have tended to focus on particular levels of representation, rather than on building a fully comprehensive model.

For example, Goldinger (1996, 1998) argues that an exemplar approach to speech perception is supported by evidence of the facilitatory effect of familiarity with a voice on recognition and memory tasks. However, his exemplar model is based on word-level episodes that by definition exclude phonetic information (potentially including voice- or talker-specific information) that relates to higher levels of linguistic structure.

Other recent exemplar approaches to speech perception focus on the importance of non-linguistic contexts beyond the individual talker, such as social group, gender and age (Docherty et al., 2006; Docherty, 2007; Foulkes & Docherty, 2006; Hay et al., 2006a; Hay et al., 2006b; Johnson, 2006). In contrast, the role of linguistic structures is emphasised in Pierrehumbert's (2003) exemplar account: 'positional variants are the workhorses of phonetic encoding' (p. 130).

Polysp, which stands for 'polysystemic speech perception' (Hawkins & Smith, 2001), is a largely episodic approach that is particularly comprehensive and explicit in terms of the structures that are hypothesised to be relevant to speech perception. Systematic phonetic detail is assumed to inform the listener simultaneously about segments, prosodic hierarchy, morphology, grammatical function, affective contexts, discourse structure and talker-related contexts.

The different emphases of these exemplar approaches illustrate that, in order to develop a coherent and comprehensive exemplar model, it is necessary to explore the levels of representation with which listeners can to learn to associate novel phonetic detail. Failure to learn certain associations would indicate that constraints on an exemplar model are necessary. For example, a failure to learn that novel phonetic information is associated with phonological phrase boundaries might indicate that 'episodes' do not relate to linguistic structures above the level of, for example, words. In contrast, the ability to learn such an association would demonstrate that phonological phrases boundaries must be incorporated as a level of representation in a dynamic model of speech perception.

Given the perceptual relevance of phonetic information relating to many levels of representation (see Section 1.3), and the indications that context-sensitive learning can occur (Section 1.4), this thesis begins to test the prediction that listeners will be able to learn novel phonetic associations conditioned by a wide range of contexts. As is evident from the studies discussed in Section 1.4 and from Samuel and Kraljic's (2009) review of perceptual learning studies, there has been surprisingly little research in this area, and there is scope to explore perceptual learning in relation to many different contexts. The present research investigates perceptual learning of phonetic detail associated with linguistic structures above the level of the segment, because the importance of these structures has been particularly neglected in models of speech perception.

1.5.2 Constraints on learnability

At present, exemplar models are also somewhat under-specified in terms of the mechanisms governing how novel information relates to prior knowledge. There is considerable evidence for the plasticity of phonetic categories (see Section 1.4), but a degree of perceptual stability must exist because listeners are able to understand speech despite huge variability. However, the balance between plasticity and stability largely remains a mystery, and current exemplar models do not specify the constraints that govern this relationship.

A very broad conception of similarity within an exemplar-based approach (as described on p. 12) would, in principle, allow any type of (detectable) phonetic variation to become associated with any context, and, moreover, for that association to change over time. However, such a flexible and unconstrained model may not be appropriate. In order to account for the broad phonetic categories that enable listeners to equate meanings and messages across speakers, styles and rates of speech, the model must be constrained such that learning from recent experience does not over-generalise to inappropriate situations. That is, there must be constraints governing how new information is integrated with past experience.

By investigating perceptual learning in relation to a range of contexts, the present research aims to shed light on the constraints on what is learnable, and thus to help elucidate the mechanisms governing the plasticity-stability relationship.

1.6 OUTLINE OF THESIS

This thesis investigates the effect that experience of novel context-dependent phonetic characteristics has on the perception of adult listeners, aiming to provide evidence about the contextual information that is encoded in phonetic representations on an ongoing basis, and whether there are constraints on the phonetic associations that are learnable.

The focus of this research is perceptual learning relating to linguistic-structural contexts. It uses relatively natural stimuli and tasks to investigate perceptual learning of phonetic detail relating to multiple levels of representation, specifically syllable stress (Chapter 2), grammatical class (Chapter 3) and morphology (Chapter 4). Results are discussed in the context of proposed properties of an exemplar model of speech perception in Chapter 5.

CHAPTER 2

EXPERIMENT 1: PERCEPTUAL LEARNING OF A STRESS-CONDITIONED ALLOPHONE

2.1 BACKGROUND

As described in Chapter 1, this thesis aims to elucidate a dynamic, context-sensitive model of speech perception by investigating the extent to which listeners can learn to associate novel phonetic cues with higher levels of linguistic structure. Listeners have been shown to use subtle phonetic cues relating to multiple levels of representation to facilitate perception (see Section 1.3), and, despite the scarcity of research in this area, there is some evidence that listeners can adapt to novel context-sensitive phonetic information (see Section 1.4). Therefore, it seems likely that listeners should be able to learn associations between phonetic cues and multiple levels of representation.

To begin to test this hypothesis, Experiment 1 investigated perceptual learning of a change in vowel quality conditioned by syllable stress. Syllable stress was selected as the conditioning context because stress has strong influence on the phonetic realisation of segments, and syllable stress is a powerful perceptual cue (e.g., Cooper et al., 2002; Cutler & Norris, 1988; van Donselaar et al., 2005). It seems probable, therefore, that listeners should be able to learn novel phonetic association with syllable stress.

Before describing the details of Experiment 1, the 'naturalistic' approach adopted by this research and its methodological implications are discussed.

2.1.1 Perceptual learning: The importance of natural stimuli and tasks

As noted in Section 1.1, the emphasis of this research is on investigating the phonetic representations that underlie *everyday* speech perception. Therefore, this thesis aims to explore methods for inducing and assessing context-sensitive perceptual learning in a relatively natural listening situation. This contrasts with many of the studies cited in Chapter 1, which frequently use synthetic stimuli and/or unnatural tasks. Sections 2.1.1.1–2.1.1.5 discuss aspects of experimental stimuli and tasks that have been shown to affect

what listeners learn from exposure to speech, and their implications for the present experiments.

2.1.1.1 Synthetic speech

Synthetic speech can be used to produce tightly-controlled stimuli that enable investigation of the perceptual influence of particular phonetic cues. However, the extent to which results can be generalised to listeners' behaviour when listening to natural speech is debatable.

Firstly, cues which are perceptually useful in synthetic speech may not necessarily be useful in natural speech. For example, Strange and Dittmann (1984) used a same-different discrimination task with immediate feedback to train eight Japanese listeners on a synthetic *rock–lock* continuum. Training improved performance in identification and oddity discrimination tests for seven subjects for the *rock–lock* series and generalised to a synthetic *rake–lake* continuum for five subjects. However, there was no significant generalisation of training to a minimal pair discrimination task using 16 pairs of words of natural speech.

The complexity of the stimulus may also influence which cues are used by the listener. Miller and Liberman (1979) found that listeners compensated for speech rate in response to a synthetic /ba/–/wa/ continuum; listeners' /b/–/w/ phoneme boundary moved towards longer transition durations as the duration of /a/ increased. However, the syllable duration effect was eliminated when a 'more natural' /ba/–/wa/ continuum was used, altering not only transition duration, but also formant trajectories and the onset frequencies of formant transitions (Shinn et al., 1985). This suggests that a phonetic characteristic may play one role in the context of the more restricted cues available from synthetic speech, but have a different effect within the more complex acoustic characteristics of natural speech. So, although generalisation from synthetic training to a natural test environment can occur (e.g., Morosan & Jamieson, 1989), caution is necessary before drawing conclusions about natural speech on the basis of the role of a phonetic cue in synthetic speech.

Some studies using synthetic speech have used exaggerated cues in order to investigate adaptation. For example, Allen and Miller (2004; see Section 1.4.8) used an extended VOT of 170–180 ms for their long-VOT talker. This is at the extreme end of the normal English voiceless VOT range and therefore their conclusion, that listeners are sensitive to individual talker differences in VOT and that this may benefit perception, could be restricted to the

very particular circumstances of their experiment and may have limited relevance for adaptation in most everyday contexts.

Pisoni (1997) describes synthetic speech as 'an impoverished signal that represents phonetic distinctions with only the minimum number of acoustic cues used to convey contrasts in natural speech' (p. 554). As part of a series of perceptual experiments, he found that synthetic speech is less robust in noise and more difficult to recall than natural speech, and concluded that synthetic speech requires more processing capacity and greater short-term memory resources than natural speech. Such processing differences could influence which aspects of speech are encoded, and thus affect perceptual learning.

This thesis specifically aims to explore the phonetic characteristics listeners are likely to learn about in everyday conversation, and so the use of relatively natural speech stimuli was prioritised over absolute phonetic control of the stimuli.

2.1.1.2 Encoding of relevant phonetic detail: The role of variability

The polysystemic, context-sensitive approach that is adopted in this thesis emphasises the importance of all kinds of context to speech perception, whether linguistic or relating to pragmatic or social situations. However, the relevance of a particular context or a specific phonetic cue in a particular listening situation might depend on many factors.

Training with a stimulus does not automatically lead the listener to encode *all* aspects of that stimulus into a phonetic representation. Morosan and Jamieson (1989; see also Section 1.4.4) trained monolingual Canadian francophones to identify the syllables /θʌ/ and /ðʌ/. Although training using synthetic CV stimuli improved identification of /θʌ/ and /ðʌ/ from natural tokens of CV stimuli from four different talkers, the training did not improve subjects' performance in distinguishing /ðʌ/ from /dʌ/ for either natural or synthetic tokens. The authors concluded that 'frication' was not encoded into the category representation for /ð/. The training was presumably insufficient for listeners to determine all of the relevant phonetic characteristics that contribute to the /ð/ category, because listeners learned to focus only on the features distinguishing /θʌ/ from /ðʌ/. This shows that even strong acoustic cues that are available in the stimulus may not be encoded if they are not deemed relevant to the training task. In addition, a native English speaker's /ð/ category is likely to be rather different from the synthesised fricatives used in this experiment: syllable-initial /ð/ occurs almost exclusively in function words in English, and it is therefore subject to connected speech processes that render it very different to a canonical /ð/, such that it is likely to be

highly reduced and there may be little or no frication (Local, 2003; Manuel, 1995). Therefore, a representation of /ð/ that would generalise effectively to assist perception of natural speech was unlikely to be learned in this study.

Introducing more variability in training can help listeners to more accurately identify the phonetic characteristics that are most relevant to particular categories. In a study of Japanese learners' acquisition of the English /l/–/r/ distinction, Lively et al. (1993) found that training with multiple talkers generalised to tokens of novel words produced by a familiar talker and by an unfamiliar talker. In contrast, training with a single talker produced stimulus-specific learning that failed to generalise to either novel words or novel talkers. It is likely that exposure to a wide range of tokens enables the listener to identify the phonetic characteristics that are relevant to the phonological category being learned, leading to more robust phonological categories that will generalise across contexts.

The importance of variability in creating more robust categories is also evident in studies that investigate perceptual adaptation to non-native accents. Bradlow and Bent (2007) used a sentence transcription task to familiarise American listeners with Chinese-accented English. Training with a single talker improved performance for that talker but failed to generalise to novel Chinese-accented talkers, whereas training with Chinese-accented multiple talkers did generalise to novel Chinese-accented talkers.

In addition to talker variability, linguistic and stylistic variability can also contribute to the formation of robust phonological categories. In contrast to Bradlow and Bent's study, Weil (2001) found that training American listeners with stimuli from one Marathi-accented talker generalised to facilitate perception of a novel Marathi-accented talker. However, although exposed to only a single talker, listeners' training with the Marathi-accented talker was both longer and more varied than in the Bradlow and Bent study – it included word identification, anomalous and meaningful sentence transcription and comprehension of short prose passages.

The range of linguistic structures used in training is also important because listeners are known to use phonological, lexical, semantic and syntactic knowledge to guide their interpretation of the talker, and, unsurprisingly, these sources of information facilitate perceptual learning (Cutler et al., 2008; Norris et al., 2003; Bradlow & Bent, 2007). Therefore, the degree to which listeners are able to make use of these kinds of linguistic knowledge during training will affect what is learned.

In summary, when listeners are exposed to only a narrow range of exemplars, perceptual learning appears to be relatively detailed and stimulus-specific. Exposure to a more variable set of exemplars (in terms of talkers, linguistic contexts and styles) enables listeners to determine which cues are relevant in a particular situation or task, making possible the formation of broader categories such as allophones or phonemes.

In natural speech, a particular phoneme or word may occur in many different phonetic, prosodic and possibly stylistic contexts, and even the same speaker uttering the same phrase will be acoustically variable. The present research aimed to explore what can be learned from exposure to speech in a fairly ordinary listening situation, so training did not include any repeated (acoustically-identical) tokens, and the phonetic characteristic under investigation appeared in a range of linguistic contexts (with the exception of the conditioning context, which was held constant). For simplicity, training in each experiment involved a single talker using a reasonably formal speech style.

2.1.1.3 Encoding of relevant phonetic detail: The role of attention

The effectiveness of training is affected not only by the stimuli, but also by the task, which can act to attract attention to or divert attention from particular aspects of the stimuli. Loebach et al. (2008) varied the task that subjects had to perform during training with vocoded (spectrally-degraded) speech. All subjects heard the same stimuli but either had to perform gender identification of the talker, talker identification or sentence transcription. In a subsequent sentence transcription task, subjects from the talker identification and transcription training groups performed significantly better than subjects from the gender identification training group. The authors conclude that training tasks that required attention to fine phonetic detail led to better performance in extracting linguistic information from the speech signal.

The importance of task in speech processing is further emphasised by neuroscientific data. For example, Hickok and Poeppel (2004) cite evidence which suggests that tasks that involve explicit access of sublexical structures (e.g., phoneme monitoring) activate temporal–parietal systems (the 'dorsal stream'), while tasks that involve mapping sounds onto meaning activate the cortex around the junction of the temporal, parietal and occipital lobes (the 'ventral stream').

Many of the studies of perceptual learning mentioned in Chapter 1 trained listeners using tasks that differ considerably from an ordinary listening situation – for example, lexical

decision tasks (Norris et al., 2003, and later studies using the same paradigm), phrase-picture matching (Peperkamp and Dupoux, 2007) and phoneme or syllable categorisation with feedback (Allen & Miller, 2004; Francis et al., 2000; Morosan & Jamieson, 1989). Such tasks will direct attention to particular aspects of the speech signal, and so affect what is learned. Furthermore, the use of isolated words, nonsense syllables and artificial languages renders 'listening' in these studies a very different task to understanding connected speech: word segmentation may not be needed, prosodic cues will be hugely different (e.g., list intonation vs. intonational phrases), connected speech processes will differ (e.g., due to the presence vs. absence of function words) and semantic access may not be required. Again, this is likely to affect the attention paid to particular aspects of the stimuli.

In order to maximise applicability to everyday listening situations, the training given to listeners in the present research used connected speech, and involved listeners trying to understand the meaning of what they hear, rather than performing a meta-linguistic task.

2.1.1.4 The importance of the assessment task

The task used to assess what has been learned from training is also crucial, because different tasks can lead to different conclusions about what has been encoded. For example, Bradlow et al. (1999) found that presenting repeated words at the same amplitude in a recognition memory task did not improve listeners' performance, although there was an advantage for words spoken by the same talker or at the same rate. Nevertheless, listeners could explicitly report whether a word was repeated at the same amplitude as a previous token. Thus, phonetic information may be retained without being brought to bear on a particular task. This conclusion is supported by the results of Weil (2001; see Section 2.1.1.2): subjects generalised learning to a novel talker when assessed in a sentence transcription task, but an (isolated) word transcription task did not reveal generalisation of learning.

The distinction between what is encoded and what is retrieved in a particular situation is well-known and has been extensively studied within psychology (e.g., Deese & Hulse, 1967, p. 379; Logan et al., 1996). In methodological terms, it highlights the importance of the assessment task in revealing what has been learned from training. When elucidating models of speech perception, it is important to investigate not only the types of phonetic information that are encoded in memory, but whether (or under what circumstances) they are perceptually useful. As the present research aimed to investigate the extent to which

newly-encoded information is used in normal speech perception, the assessment tasks were designed to resemble a natural listening situation as far as possible.

2.1.1.5 Summary of the naturalistic approach

Chapter 1 emphasised the importance of conceptualising the processing of the speech signal in terms of all relevant contextual information. This approach receives further support from evidence that stimulus complexity, attention, and stimulus variability influence what is learned by the listener. To achieve a realistic idea of the phonetic cues that listeners might adapt to when they hear a particular talker in an ordinary listening situation, it is necessary to replicate that ordinary listening situation as far as possible; this includes the characteristics of the training stimuli, and the task during training and assessment.

2.1.2 Methodological background to Experiment 1

Experiment 1 investigated whether listeners can learn novel associations between phonetic cues and prosodic structure by studying perceptual learning of a stress-conditioned allophone. Specifically, it investigated whether listeners can adapt to a novel accent in which /ɪ/ is mid-centralised (or 'reduced') to [ə] in unstressed syllables. (Henceforth, 'mid-centralisation' will be referred to as 'centralisation'.)

Centralisation of /ɪ/ to [ə] was selected for several reasons. Firstly, it is (in some words at least) a phonemic distinction (e.g., *Lenin*, *Lennon*) in British English, and therefore should be a salient acoustic change for British English listeners.

Secondly, it was desirable to use a natural-sounding variant, but one that is nevertheless an atypical accent feature and might therefore induce measurable perceptual learning. [ə] is a natural-sounding variant of /ɪ/ in unstressed syllables – it distinguishes some accents; for example, the suffix *-es* is generally pronounced [ɪz] in British English but [əz] in US English (Fabricius, 2002). There is also a sound change in progress in British English whereby [ə] is becoming increasingly prevalent in unstressed syllables previously pronounced as [ɪ] (Fabricius, 2002; Wells, 1999). However, there are still many words which are consistently pronounced as [ɪ] in unstressed position (e.g., *public*, *pulpit*, *attic*, *illegal*) by the majority of speakers, so realising /ɪ/ as [ə] in all unstressed syllables should be an unusual accent feature for British listeners.

Thirdly, /ɪ/ (realised as [ɪ]) occurs in stressed and unstressed syllables in British English. This enables a comparison between perceptual learning of an accent in which /ɪ/ is centralised to [ə] in unstressed syllables only, and an accent in which /ɪ/ is centralised in

both stressed and unstressed syllables. Centralisation of /ɪ/ in stressed and unstressed syllables occurs in New Zealand English as well as some Scottish accents, and therefore it should not sound unnatural to British English listeners because they are likely to have heard it at least occasionally on broadcast media.

To create a fairly natural listening environment, the training task[1] in Experiment 1, whereby subjects were familiarised with the accent in question, involved listening to a story spoken by a real talker. A comprehension task was also included during training to ensure that subjects paid attention to the story and listened to it with the aim of understanding the meaning. Two previous studies have successfully used stories to familiarise listeners with an atypical accent, although Maye et al. (2008) used a MacinTalk speech synthesiser, while Eisner and McQueen (2006) created their story by splicing resynthesised ambiguous segments into natural speech.

Perceptual learning was assessed using an auditory lexical decision task containing instances of /ɪ/ realised as [ə]. Although a lexical decision task does not replicate what occurs during a normal interaction, it does involve lexical access and therefore it should detect whether experience of the talker influences listeners' interpretation of a word. This method was successful in detecting adaptation to lowered front vowels in the study by Maye et al. (2008).

2.2 METHOD

2.2.1 Subjects

62 subjects (24 male, 38 female) participated in the experiment. Subjects were non-phonetically-trained students from the University of Cambridge, aged between 18 and 30, native speakers of British English, monolingual (until at least age 11), and had no known speech or hearing problems. They were unfamiliar with the voice of the talker in the experiment, and were paid £6 for their participation. Two subjects were excluded from the final data analysis (see Section 2.4.1), leaving 60 subjects (20 subjects in each of three Training groups) in the statistical analysis.

[1] The term 'training task' rather than 'familiarisation' is used in this thesis to avoid confusion, because familiarisation with an accent can also occur during exposure to a variant in the test task. However, it is worth highlighting that the 'training tasks' do not include explicit feedback.

2.2.2 Materials

2.2.2.1 Training materials: Story

The training material was an original story read by a 23-year-old female who is a phonetically-trained native English speaker with a mild Leeds accent. She had practised reading the stories in advance to ensure fluency. Recording took place in a sound-treated room at a 22050 Hz sampling rate using Media Recorder software on a Silicon Graphics computer. A Sennheiser MKH 40 P48 high-frequency cardioid condenser microphone was used.

Three versions of the story were recorded in different 'accents': a **Global** version, a **Stress-conditioned** version and a **Control** version, as described in Table 1.

	Realisation of /ɪ/ in a stressed syllable		Realisation of /ɪ/ in an unstressed syllable		Sample sentence	Story duration
	Vowel quality	**No. of tokens**	**Vowel quality**	**No. of tokens**		
Global	[ə]	260	[ə]	260	Dame [ə]sabel was not merely an adept broadcaster who had an eloquent and clear writ[ə]ng style, but was also a woman of great bravery.	26 m 34 s
Stress-conditioned	[ɪ]	260	[ə]	520	Dame [ɪ]sabel was not merely an accompl[ə]shed broadcaster who had a flu[ə]d and luc[ə]d writ[ə]ng style, but was also a woman of great cour[ə]ge.	26 m 46 s
Control	[ɪ]	260	[ɪ]	520	Dame [ɪ]sabel was not merely an accompl[ɪ]shed broadcaster who had a flu[ɪ]d and luc[ɪ]d writ[ɪ]ng style, but was also a woman of great cour[ɪ]ge.	27 m 23 s

Table 1. Number of tokens of /ɪ/ and vowel quality of /ɪ/ produced by the talker in Global, Stress-conditioned and Control versions of the story.

The convention of not allowing [ə] in stressed syllables in RP is not followed here because the production target in the Global story was a stressed mid-central vowel with the auditory percept /ə/, approximating to the New Zealand pronunciation of /ɪ/, which is sometimes transcribed as [ə] (e.g., Giegerich, 1992; Scharinger & Lahiri, 2010; Wells, 1982).

The acoustic properties of schwa can vary considerably depending on the surrounding phonetic context, and therefore [ə] was defined in impressionistic auditory terms. The recordings were checked auditorily by the author, and any realisations that were not sufficiently clear tokens of [ə] were re-recorded.

Differences between the Global, Stress-conditioned and Control stories were minimised by ensuring (as far as possible) that each version of the story was spoken at the same impressionistic rate and with the same intonation pattern. Global and Stress-conditioned versions contained the same number of tokens of centralised /ɪ/. This was achieved by replacing some words or phrases which did not contain /ɪ/ in the Global version with words or phrases which contained unstressed /ɪ/ in the Stress-conditioned version. For example, *peaceful* (Global version) was replaced by *tranquil* (Stress-conditioned version), and *small house* (Global version) was replaced by *cottage* (Stress-conditioned version). The meaning was kept as similar as possible between the two versions. In both versions of the story, 30% of the tokens of centralised /ɪ/ occurred in affixes. Syllables containing /ɪ/ whose status as stem or affix was ambiguous (e.g., *indifferent*, *oasis*) were avoided. Control subjects heard the same story as the Stress-conditioned subjects (but with /ɪ/ pronounced as [ɪ] throughout). Both versions of the training story are in Appendix A.

Some British English speakers routinely centralise /ɪ/ in unstressed syllables in some words. The incidence of such words was minimised by using only words for which the 1997 edition of *The English Pronouncing Dictionary* had no alternative [ə] pronunciation of /ɪ/. As an extra precaution, the author listened to four speakers of British English (aged 22–30) reading the story, and any words in which /ɪ/ was pronounced as [ə] by any one of them were replaced. These criteria excluded many words, including all words ending *–ity*, *-icate*, *-able*. Function words containing /ɪ/

(e.g., *with*, auxiliary *is*) were also excluded because function words are frequently very reduced in connected speech.

Words in which /ɪ/ carries secondary stress (e.g., ˌ*libeˈration*) were excluded because including secondary stress might add complexity to the learning task and affect the learnability of the 'rule' governing [ɪ]-[ə] variation.

Words containing *–ied* (e.g., *carried*) or *–ies* (e.g., *fairies*), which are sometimes pronounced as [ɪ], were included, but were pronounced with the alternative [i:] realisation.

20 comprehension questions were included in the training task. Questions were designed to be hard enough that subjects would have to pay attention to the meaning of the story, but easy enough that they were not trying to memorise the story as they heard it. Comprehension questions are listed in Appendix B.

2.2.2.2 Test materials: Auditory lexical decision

Perceptual learning was assessed using an auditory lexical decision task. The talker and recording equipment were identical to the training task.

The auditory lexical decision task comprised 320 stimuli: 160 non-words, 80 filler words, and 80 experimental words. 40 experimental words contained /ɪ/ in a stressed syllable (e.g., ˈ*impulse*) and are referred to as **Stressed-vowel** stimuli; and 40 contained /ɪ/ in an unstressed syllable (e.g., *imˈplore*) and are termed **Unstressed-vowel** stimuli. All instances of /ɪ/ were centralised to [ə] by the talker. Stimuli are listed in Appendix C.

All lexical decision stimuli were novel (i.e. they had not occurred in training). The experimental words were selected according to the same criteria as the words used in training, i.e. the pronunciation of /ɪ/ is listed as [ɪ] in the 1997 *English Pronouncing Dictionary* and pronounced as [ɪ] by all four speakers of British English who were tested. Words in which a centralised pronunciation of /ɪ/ would create a different word were also excluded because, for example, a Word response to the word *satin* pronounced as [sætən] could indicate recognition of the word *Saturn*.

To prevent an advantage for one subject group due to prior exposure to particular phoneme sequences, the phoneme sequences were matched across Stressed-vowel and

Unstressed-vowel stimuli. This was necessary because Global and Stress-conditioned subjects were exposed to /ɪ/ in different phonemic contexts in training; for example, Global subjects had frequent exposure to [əz] in the word *Isabel*. Stressed-vowel and Unstressed-vowel stimuli were therefore selected in pairs, such that the same phoneme sequence occurred in the key syllable for each pair e.g., ˈ*disco*–ˈ*bodice*, *con*ˈ*vince*–ˈ*province*.

The number of eligible Stressed-vowel and Unstressed-vowel stimulus pairs is quite restricted, so proper names were included in the task, and subjects were instructed that names should be treated as real words. For the same reason, lexical frequency and position of the key syllable were not controlled across Stressed-vowel and Unstressed-vowel stimuli.

In order to encourage rapid 'instinctive' decisions, non-words had a sparse lexical neighbourhood. Words and non-words were matched for syllable number and stress pattern to prevent these factors being predictive of lexical status.

2.2.3 Procedure

Subjects were tested individually in a sound-treated room and used high-quality headphones. All stimuli were played from a desktop PC running Windows XP using DMDX software (Forster & Forster, 2003). Subjects were given written instructions about the experiment, and the experimenter answered any questions the subjects had.

The experiment began with a practice session. This comprised a short prose passage, followed by an auditory lexical decision task (10 stimuli) and a comprehension question. Stimuli were spoken by the same talker as was heard in the main session and did not contain the phoneme /ɪ/ (see Appendix D for practice materials).

Training and testing tasks were interleaved throughout the experiment in order to keep the test tasks short, thus minimising the possibility of subjects adapting to the atypical pronunciation during the test task.

The training story was divided into 20 sections, with an average section length of 80 seconds (Global), 80 seconds (Stress-conditioned) and 82 seconds (Control). Subjects were requested to listen carefully to the story, bearing in mind that they would be asked questions about it afterwards. Subjects were randomly assigned to one of three

Training groups, and heard only one version of the story (Global, Stress-conditioned or Control) during training.

After each section of story, subjects completed an auditory lexical decision task comprising 16 stimuli. The order of the stimuli was controlled such that each word within a Stressed-Unstressed pair occurred in a different half of the experiment. Within this constraint, filler words, experimental words and non-words were presented in random order. The order was counterbalanced across pairs of subjects within each Training group. The ISI was 3700 ms, leaving subjects with an average of 2864 ms to respond after word offset. Subjects responded by pressing the left or right buttons on a mouse connected to the PC via a serial port. Subjects used left and right index fingers to respond. The mouse was placed 'back-to-front' and attached to the table in order to make the subjects' hand position more comfortable. Responses ('YES' for words or 'NO' for non-words) were presented on either the left or the right of the screen as the stimulus was heard. The assignment of positive and negative response buttons was counterbalanced within each condition. Subjects were instructed to respond as quickly and accurately as possible, and response times were recorded. (Response times recorded by DMDX using a serial mouse are reported to have a standard deviation of only 0.55 ms from the actual response time (Forster & Forster, 2003)).

One comprehension question was presented after each auditory lexical decision task. The question was displayed at the top of the screen, and two possible answers were displayed on the left and right of the monitor. Subjects responded by pressing the left or right mouse button as appropriate. Again, the assignment of correct response to the left or right mouse button was counterbalanced across subjects within each condition.

This sequence (story, auditory lexical decision task, comprehension question) was repeated 20 times. The experiment took approximately one hour to complete.

2.3 HYPOTHESES

Performance in the test task was predicted to reflect experience during training: stimuli that were pronounced consistently with the accent heard during training were expected to have a higher lexical acceptance rate. Therefore, Global subjects were predicted to have a higher lexical acceptance rate than Control or Stress-conditioned

subjects for Stressed-vowel stimuli. For Unstressed-vowel stimuli, both Global and Stress-conditioned subjects should have a higher lexical acceptance rate than Control subjects.

Training and test tasks were interleaved 20 times throughout this experiment. To assess the rate of learning, the variable Task number (1 to 20) was included in the statistical analysis. Little is known about the rate at which subjects can learn about phonetic detail conditioned by linguistic context, particularly from continuous speech. Therefore, it was possible that all learning could occur within the first training task, giving no effect of Task number. Alternatively, if learning occurred at a slightly slower rate, Global and Stress-conditioned subjects' responses were predicted to become more consistent with the accent heard during training as Task number increased. Perceptual learning from exposure to the test stimuli could also be a factor, leading to greater acceptance of /ɪ/ as [ə] for Control subjects as Task number increased.

It is uncertain how far a stimulus has to deviate from the standard pronunciation of a word before subjects classify it as a non-word. Consequently, it was possible that both Stressed-vowel and Unstressed-vowel stimuli would be classified as acceptable words in the context of the experiment, in which case differences were predicted to be reflected in response times. Response times were predicted to be faster for responses that were consistent with the accent heard during training.

2.4 RESULTS AND DISCUSSION

2.4.1 Data pre-processing

Two participants were excluded from the analysis: one male Global subject had a 5% 'no response' rate in the lexical decision task and one female Stress-conditioned subject had a 25% error rate for filler non-words (compared with an average 1% error rate for other participants). There were 20 subjects in each Training group in the statistical analysis.

Trials for which no response was produced before the deadline accounted for 0.0006% of the Stressed-vowel and Unstressed-vowel trials and were excluded from the analysis.

An average of 97% of the comprehension questions were answered correctly, with subjects ranging from 90–100% correct.

2.4.2 Lexical decision task: Lexical acceptance rates

2.4.2.1 Statistical technique for binary responses

Experiments 1 and 2 in this thesis were analysed in R version 2.12.0 (R Development Core Team, 2010) using logistic regression. Although binary data is frequently analysed using ANOVA, logistic regression is recommended (e.g., Baayen, 2004) because it assumes that response data are binomially distributed and so it can be used to fit a model using a data set that includes every individual binary response. This contrasts with ANOVA, which use a normally-distributed data set created by calculating average percentages for each subject or each item. Thus, using logistic regression avoids the loss of information through averaging. In addition, logistic regression is preferred to linear regression because it prevents predicting unrealistic probabilities that are less than 0 or more than 1.

Data were analysed using generalised linear mixed-effects modelling (as described in Baayen, 2007). Mixed-effects models take into account variance from both fixed and random factors, thus increasing statistical power. Backwards regression was used to identify the most parsimonious model. That is, a full model with all fixed and random factors and all interactions was fitted to the data, and predictors that did not improve the fit of the model were incrementally removed until the most parsimonious model was found. The significance of predictors was assessed using log likelihood tests. Backwards regression was used because this work is relatively exploratory, particularly in relation to the rate of perceptual learning (see Section 2.3).

2.4.2.2 Analyses of lexical acceptance rates

Data for Stressed-vowel and Unstressed-vowel stimuli were analysed separately using mixed models with fixed predictor variables of Training group (Global, Stress-conditioned, Control), Task number (1 to 20) and an interaction between Training group and Task number, and random factors of Subject and Item.

Stressed-vowel stimuli

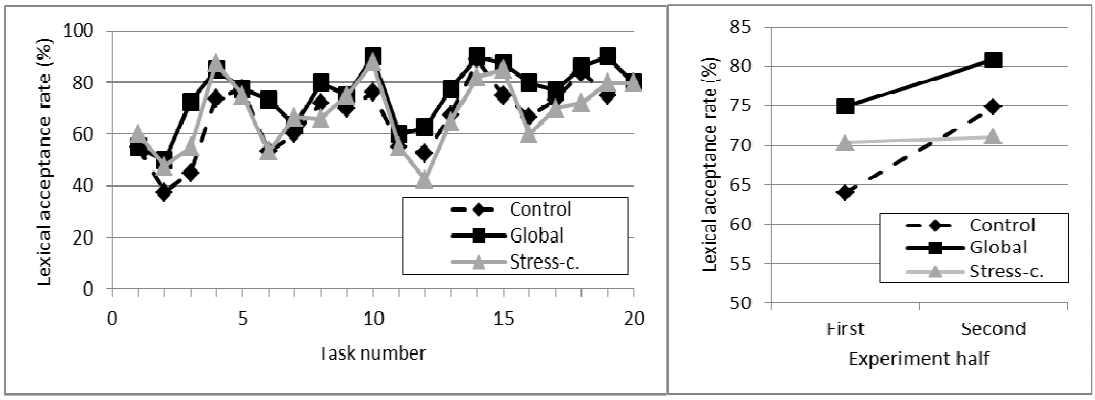


Figure 1. Mean lexical acceptance rate for Stressed-vowel stimuli in the lexical decision task, as a function of Training group and Task number (left panel), and as a function of Training Group and Half (right panel).

Results for Stressed-vowel stimuli are presented in Figure 1. The most parsimonious model using the data from Stressed-vowel stimuli included an interaction between Training group and Task number ($\chi^2 = 6.7$, $df = 2$, $p < 0.04$). The effect of Task number was significant for Global subjects ($z = 2.6$, $p < 0.01$) and Control subjects ($z = 4.5$, $p < 0.001$), but not for Stress-conditioned subjects ($z = 1.0$, $p > 0.3$).

There were insufficient data to test the effect of Training Group within each individual task. Therefore, to investigate how the effect of Training Group changed between the beginning and end of the experiment, a mixed-effects model was fitted to data from the first half (Tasks 1 to 10) and second half (Tasks 11 to 20) of the experiment. The model had fixed factors of Training group (Global, Stress-conditioned, Control), Half (First, Second), and Training group × Half, and random factors of Subject and Item. Results showed the same effects of learning as the Task number analysis: there was a significant interaction between Training group and Half ($\chi^2 = 7.8$, $df = 2$, $p < 0.03$), and there was a significantly higher lexical acceptance rate in the second half than the first half for Global subjects (75% in the first half vs. 81% in the second half, $z = 2.6$, $p < 0.01$) and Control subjects (64% vs. 75%, $z = 4.4$, $p < 0.001$), but not for Stress-conditioned subjects (70% vs. 71%, $z = 0.4$, $p = 0.7$). Pairwise comparisons showed that Global subjects had a significantly higher lexical acceptance rate than Control subjects in the first half (75% vs. 64%, $z = 2.2$, $p < 0.03$), but not the second half (81% vs. 75%, $z = 1.5$, $p = 0.1$). In contrast, there was no significant difference between the lexical acceptance rates for Global and Stress-conditioned subjects in the first half (75% vs. 70%, $z = 1.1$, $p = 0.3$), but Global

subjects had a higher lexical acceptance rate than Stress-conditioned subjects in the second half (81% vs. 71%, $z = 2.2$, $p < 0.03$). There was no significant difference between the lexical acceptance rates for Stress-conditioned and Control subjects in either half.

Unstressed-vowel stimuli

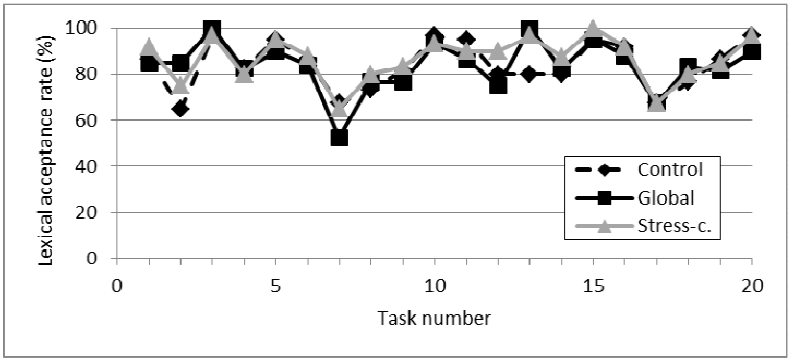


Figure 2. Mean lexical acceptance rate for Unstressed-vowel stimuli in the lexical decision task, as a function of Training group and Task number.

Lexical acceptance data for Unstressed-vowel stimuli are presented in Figure 2. A mixed-effects model showed a significant effect of Task number ($\chi^2 = 5.3$, $df = 1$, $p < 0.03$), with a higher lexical acceptance rate as the experiment progressed, increasing from an average of 84% in the first half to 86% in the second half. Contrary to the prediction, there was no significant difference in lexical acceptance rate between Training groups for Unstressed-vowel stimuli, and no significant interaction.

2.4.3 Discussion of lexical acceptance rates

2.4.3.1 Stressed-vowel stimuli

Predictions were largely supported for Stressed-vowel stimuli. Global subjects had a higher lexical acceptance rate than Control subjects (in the first half of the experiment) and Stress-conditioned subjects (in the second half of the experiment). Furthermore, Global subjects' responses became more consistent with the accent heard during training as the experiment progressed, as shown by their higher lexical acceptance rate for Stressed-vowel stimuli as Task number increased. The lack of an effect of Task number for Stress-conditioned subjects is also consistent with their training – Stress-conditioned subjects were not exposed to /ɪ/ as [ə] in stressed syllables in the story.

However, Control subjects adapted to the atypical pronunciation in Stressed-vowel stimuli despite not hearing this pronunciation during training. This adaptation was slower than that of Global subjects, as indicated by Control subjects' lower lexical acceptance rates in the first half of the experiment, but was nevertheless highly significant.

It is likely that Control subjects' adaptation was caused by exposure to the atypical pronunciations during the lexical decision task. This explanation receives support from Norris et al. (2003, see Section 1.4.2) and subsequent studies using the same paradigm; a lexical decision task was used to familiarise subjects with atypical pronunciations, and this successfully induced adaptation (as assessed by a phoneme categorisation task). Furthermore, adaptation to a new accent can be rapid; Clarke and Garrett (2004) demonstrated adaptation to non-native accents within one minute of exposure. In the present experiment, training and assessment tasks were interleaved in order to minimise the possibility of adaptation during the assessment task, but the results indicate that this was not successful. The slower rate of learning for Control subjects than Global subjects is unsurprising because Control subjects heard the atypical pronunciations only in the lexical decision task, while Global subjects could also learn from training.

It is interesting to consider why Control subjects increased their acceptance of Stressed-vowel stimuli over the course of the experiment while Stress-conditioned subjects did not show any learning effect. Neither group were exposed to [ɪ] realised as [ə] in stressed syllables during training, and both were exposed to them during the lexical decision task.

It is possible that Control subjects dissociated the training and assessment tasks because the 'accent difference' between the training and assessment tasks was greater for Control subjects than for Stress-conditioned subjects. That is, despite the same talker being used throughout the experiment, the atypical pronunciations in the assessment task may have rendered the training irrelevant for Control subjects. In contrast, for Stress-conditioned subjects, the atypical pronunciations in unstressed syllables during training were sufficient to maintain the link between training and assessment. This suggests that, where the 'accent difference' is big enough, the

influence of lexical knowledge in the lexical decision task can override knowledge of the talker's accent as heard during training.

2.4.3.2 Unstressed-vowel stimuli

Contrary to the hypothesis, Training group did not affect responses to Unstressed-vowel stimuli. The increasing acceptance of Unstressed-vowel stimuli by all Training groups over the course of the experiment may indicate a degree of perceptual learning. However, it could also be evidence of improvement at the lexical decision task itself. The latter analysis is supported by an analysis of filler non-words. Responses to filler non-words showed a significant effect of Task number ($\chi^2 = 7.9$, $df = 1$, $p < 0.01$), with more accurate responses as Task number increased (1.6% 'Word' responses in Tasks 1 to 10 vs. 1.0% in Tasks 11 to 20). Filler words do not show an effect of Task number, but this may be due to a ceiling effect. If the small increase in lexical acceptance rate for Unstressed-vowel stimuli over the course of the experiment is assumed to be due to task learning, there is no evidence of adaptation to the atypical pronunciation in unstressed syllables.

Listeners may have failed to adapt to the atypical pronunciation in unstressed syllables because the vowel centralisation is likely to have been less perceptually salient and of lower informational value in unstressed syllables than in stressed syllables (see Bond, 1981; Cole & Jakimik, 1980). There are several reasons for this. Firstly, unstressed syllables are acoustically less prominent than stressed syllables. Secondly, a wider range of vowels can occur in stressed syllables than in unstressed syllables, and therefore an atypical pronunciation in an unstressed syllable is likely to have been less disruptive to lexical access. Thirdly, syllable position may have played a role. Both within this experiment and in English more widely, /ɪ/ occurred early in the word more frequently in stressed syllables than unstressed syllables. This is likely to affect the informational value of the variant because, as described by Cohort Theory (Marslen-Wilson & Welsh, 1978), there will be fewer lexical competitors late in the word, making the identity of the segment relatively less important. Fourthly, although the words in the lexical decision task were selected to minimise prior experience of /ɪ/ being realised as [ə], the acceptability of /ɪ/ → [ə] in other words may have decreased its perceptual salience. Finally, in connected speech there may be added incentive to pay relatively more attention to stressed than unstressed syllables;

the Metrical Segmentation Strategy (Cutler and Norris, 1988) suggests that English listeners use stressed syllables as a cue to word segmentation.

The lack of adaptation to Unstressed-vowel stimuli may also be a consequence of performance that is close to ceiling in the lexical decision task. Subjects accepted 85% of Unstressed-vowel stimuli as words, which may limit the scope for any perceptual learning effects to be detected. The effective ceiling will be lower than 100% because of errors by subjects; after the experiment, several subjects reported that they had realised a stimulus was a real word after pressing the "No" (Non-word) button. The relatively rapid responses encouraged by the lexical decision task may limit the potential for atypical pronunciations to be accepted, because in some cases lexical access is not fast enough to prevent a Non-word response.

In addition, it may be difficult to induce acceptance of certain words when they are pronounced atypically, even after adaptation. Even in studies where perceptual learning has been demonstrated, the percentage of stimuli accepted as words does not approach 100%: Global subjects in this experiment showed clear evidence of perceptual learning for Stressed-vowel stimuli, but accepted only 78% of Stressed-vowel stimuli as words (and no subject accepted more than 95%). This compares to a lexical acceptance rate of 97% for word fillers for Global subjects. Similarly, in Maye et al.'s study (2008, see Section 1.4.2), those items incorporating the vowel shift were accepted on only 69% of occasions (in their Experiment 2) despite strong evidence of perceptual learning. Studying listeners' responses over a longer training period would be necessary to establish whether it is possible to achieve equivalent performance for stimuli incorporating a newly-learned pronunciation and stimuli that do not.

2.4.4 Lexical decision task: Response times

Response times were measured from word offset. Negative response times were included in the analysis, because some words may have been identified before word offset, and there were no clear outliers. Only Word responses (i.e. 'correct' responses) were included in the analysis.[2] The distribution of response times had a substantial

[2] It could be argued that separating Word and Non-word response times is inappropriate in this experiment because the 'correct' response to Stressed and Unstressed keywords is dependent on subjects' interpretation of the stimulus. For this reason, analyses which included Non-word responses were also carried out. These showed the same pattern of results as those excluding Non-word responses.

right-tail, so, in order to reduce skewness, a constant was added to response times (to ensure all were more than zero) and the adjusted response times were log-transformed.

Stressed-vowel and Unstressed-vowel stimuli were analysed separately. For each, a linear mixed-effects model was fitted with fixed predictors of Training group (Global, Stress-conditioned, Control), Task number (1 to 20) and an interaction between Task number and Training group, and random factors of Subject and Item. Predictors that did not significantly contribute to the model were removed incrementally until the most parsimonious model was found.

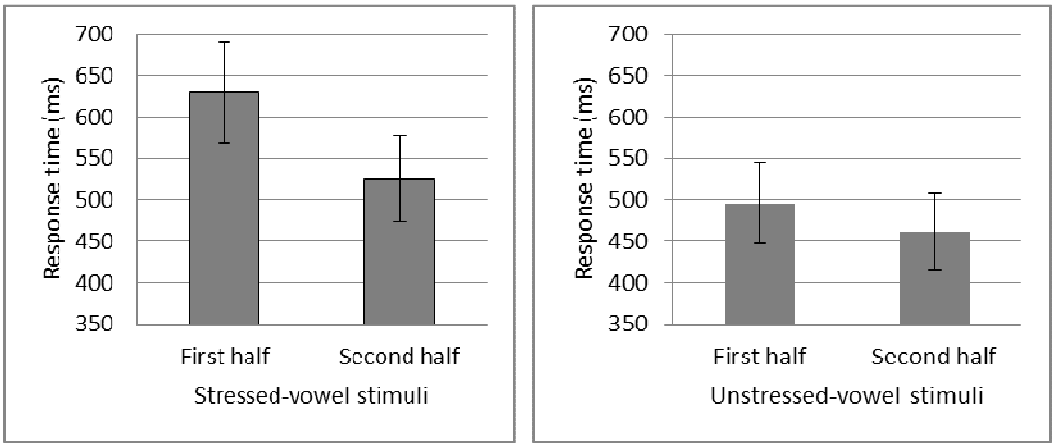


Figure 3. Mean response time (ms) for Stressed-vowel stimuli (left panel) and Unstressed-vowel stimuli (right panel) in the lexical decision task, as a function of Task number (1 to 10 vs. 11 to 20). Error bars show the 95% confidence interval around the mean.

For both Stressed-vowel and Unstressed-vowel stimuli, only Task number was a significant predictor in the most parsimonious model (Stressed-vowel: $\chi^2 = 73.2$, $df = 1$, $p < 0.001$; Unstressed-vowel: $\chi^2 = 15.4$, $df = 1$, $p < 0.001$). In both cases, responses times were shorter as Task number increased. Figure 3 plots the response times for the first half (Tasks 1 to 10) and second half (Tasks 11 to 20) of the experiment (when response times were plotted as a function of Task number, between-item variation masked the trend towards shorter response times, so this figure is not presented).

2.4.5 Discussion of response time analyses

Results of the response time analysis failed to support the hypothesis that response times should be shorter when stimuli are consistent with the accent heard during

training. There were no significant effects involving Training group within either Stressed-vowel or Unstressed-vowel stimuli.

The significant effect of Task number showed that response times became shorter as the experiment progressed. This is at least partly an effect of task practice: response times to filler words decreased as Task number increased ($\chi^2 = 7.6$, $df = 1$, $p < 0.01$) [3], from an average of 364 ms for Tasks 1 to 10 to 349 ms for Tasks 11 to 20, and response times to filler non-words also decreased with increasing Task number ($\chi^2 = 165.8$, $df = 1$, $p < 0.001$) [4], from 499 ms for Tasks 1 to 10 to 452 ms for Tasks 11 to 20.

It is notable that the perceptual learning effect for Global subjects for Stressed-vowel stimuli (indicated by their higher lexical acceptance rate) is not reflected in the response times. This raises the possibility that response times in a lexical decision task may not be an appropriate measure of perceptual learning. Maye et al. (2008) also failed to find any differences in response times that reflected the perceptual learning effect that was evident in their analysis of lexical acceptance rate in a lexical decision task.

The predicted effect of Training group on response time was based on the assumption that perceptual learning would make the task of accessing words in the lexical decision task easier and therefore faster. However, there are two situations in which this may not hold true.

Firstly, perceptual learning (at least in the context of fairly limited training) may not lead to faster processing of the atypical pronunciation.

Secondly, responses to the lexical decision task may not be sufficiently automatic. That is, subjects may respond in a strategic manner that makes response times more variable and less closely-related to task difficulty. The latter explanation appears to be true of at least some subjects. During informal questioning after the experiment, several subjects specifically mentioned adopting a strategy of either accepting

[3] A constant was added to 'Word' response times, which were then log-transformed. A mixed-model with a fixed factor, Task number, and random factors of Subject and Item was fitted.

[4] One data point was excluded as it was a clear outlier, more than 400 ms faster than any other response. A constant was added to the 'NonWord' response times, which were then log-transformed. A mixed-model with a fixed factor, Task number, and random factors of Subject and Item was fitted.

('making allowances') or rejecting words that appeared to be slightly mispronounced. The decision to use non-words which were very distinct from real words was designed to encourage 'instinctive' responses for the entire task, but it may have had the unintended effect of making the experimental words stand out, thus enabling subjects to adopt a specific acceptance or rejection strategy for these words.

Nevertheless, response times were longer for Stressed-vowel stimuli than Unstressed-vowel stimuli (an average of 578 ms vs. 478 ms), which suggests that response times reflect task difficulty to some extent. Stressed-vowel stimuli, in which centralisation was more perceptually and informationally salient, took longer to recognise as real words.

2.5 CONCLUSION

Although the results did not fully bear out the predictions, Experiment 1 nevertheless has significant theoretical implications.

Firstly, the results are consistent with those of Maye et al. (2008) and Eisner and McQueen (2006) in showing that familiarising subjects with an accent by listening to a story can produce perceptual learning. Importantly, the present study extends these results to speech which is produced by a single talker with no synthetic manipulation, and which is therefore perceptually coherent in terms of coarticulatory effects and relative segment durations. It provides evidence that perceptual learning induced through training with lexical decision tasks, synthetic or synthetically-manipulated speech and categorisation tasks may have more natural counterparts in everyday speech perception.

Secondly, although predictions were not fulfilled for Unstressed-vowel stimuli, the results from Stressed-vowel stimuli demonstrated a different pattern of learning for each Training group. These results are promising insofar as they demonstrate that there is more to perceptual learning than simply hearing an unusual realisation of a phoneme a certain number of times – the distribution of these occurrences in stressed or unstressed syllables is also important. Both Global and Stress-conditioned subjects heard 520 tokens of /ɪ/ pronounced as [ə], but Global subjects learned to accept more Stressed-vowel stimuli as real words over the course of the experiment, while Stress-conditioned subjects did not show a learning effect.

Finally, the centralisation of /ɪ/ in unstressed syllables seemed to have perceptual consequences for Stress-conditioned subjects, even though it did not affect responses to Unstressed-vowel stimuli in the lexical decision task. Hearing the atypical pronunciation in unstressed syllables may have contributed to a percept of talker-identity or talker-similarity that was sufficient to 'link' behaviour on training and assessment tasks for Stress-conditioned subjects and to limit the influence of top-down lexical feedback in the lexical decision task. This contrasts with Control subjects, who seem to dissociate the two tasks.

Overall, Experiment 1 is perhaps more informative in terms of its methodological implications for investigating context-sensitive perceptual learning than in what it can reveal about perceptual learning of a stress-conditioned allophone.

Most importantly, this study highlights the importance of assessing perceptual learning in a variety of linguistic contexts. None of the perceptual learning studies cited in Chapter 1 assessed perceptual learning in unstressed syllables, and therefore their conclusions about the locus of perceptual learning may be premature.

The results also suggest that the perceptual salience and informational value of the atypical pronunciation has a strong influence on whether that pronunciation induces measurable adaptation (see also Docherty et al., 2008), and should therefore be an important consideration when selecting the phonetic variation to be studied, and in discussing the generalisability of the results. In this experiment, one conditioning context, unstressed syllables, is intrinsically not prominent, and this may have decreased the likelihood of either inducing or measuring context-sensitive perceptual learning.

The performance of Control subjects highlights an additional consideration for perceptual learning methodology: it may not be not sufficient to use the same talker in training and assessment, and assume that listeners will complete the assessment on the basis of the training; the accent used during training and assessment may also have to be sufficiently similar.

Finally, Experiment 1 has highlighted some methodological difficulties with using a lexical decision task to assess perceptual learning. As discussed in Section 2.4.5, response times in a lexical decision task are not an effective way of assessing

perceptual learning. In addition, lexical knowledge is a strong 'driver' of perceptual learning even when the lexical decision task is kept fairly short and is interleaved with training. This makes it more difficult to detect between-group differences that result from the training materials. Furthermore, the lexical decision task is not ideal in terms of the 'naturalistic' principles outlined at the beginning of this chapter. It is a meta-linguistic task, and therefore does not reflect the way attention is directed in a normal listening situation. A task which resembles the act of comprehension more closely may provide a more realistic assessment of whether listeners use novel perceptual information in conversation.

These considerations are used to inform later experiments in this thesis. Perceptual learning of an atypical pronunciation conditioned by syllable stress is not pursued further because of the apparent difficulty in inducing measurable perceptual learning in the context of unstressed syllables, which have relatively low perceptual and informational salience. Instead, Experiments 2 and 3 explore whether grammatical and morphological contexts can influence on perceptual learning. The influence of grammatical and morphological contexts on speech perception is frequently neglected in models of speech perception, and the potential influence of these levels of representation on the adaptation of the perceptual system has scarcely been addressed in the perceptual learning literature (see Samuel & Kraljic, 2009), despite evidence that they play a role in speech perception.

Because of the success of the training story (and accompanying comprehension questions) in inducing perceptual learning in Global subjects, and its consistency with a naturalistic approach, this training task will be used throughout this thesis. However, the difficulties with using a lexical decision task to assess perceptual learning led to the use of alternative assessment tasks in Experiments 2 and 3. These assessment tasks attempted to reduce the role of feedback from higher levels of linguistic structure and avoided explicit meta-linguistic decisions.

CHAPTER 3

EXPERIMENTS 2A AND 2B: PERCEPTUAL LEARNING OF A DURATION DIFFERENCE ASSOCIATED WITH GRAMMATICAL CATEGORY

3.1 BACKGROUND

Chapter 1 described an approach to speech perception which argues that listeners are sensitive to phonetic detail that reflects multiple levels of representation within different subsystems (e.g., prosodic structure, grammatical function, emotion and social contexts), and it highlighted the paucity of research into how these levels of representation influence perception dynamically. This chapter describes two closely-related experiments designed to explore context-sensitive perceptual learning of phonetic detail that is associated with the grammatical category of a word.

3.1.1 Experiment 2a: Theoretical background – Listeners' sensitivity to phonetic detail reflecting grammatical category

As described briefly in Section 1.3.4, evidence from production and perception experiments demonstrates an association between grammatical function and phonetic properties. For example, function words can be more reduced than content words (e.g., Lavoie, 2002; Baker, 2008). In addition, there are differences in patterns of assimilation across function and content word boundaries. For example, assimilation of a final /m/ to the place of articulation of a following consonant is more likely in function words than content words; *I'm going* (where the /m/ is part of an auxiliary verb) can realised as [aĩŋ], while *the lime goes* is very unlikely to undergo such assimilation (Local, 2003).

More subtle grammatical differences can also be reflected in phonetic characteristics. Based on a corpus of telephone conversations, Jurafsky et al. (2002) describe phonetic differences between phonemically-identical words that serve a different grammatical purpose. For example, infinitival *to* tends to be shorter and more reduced than prepositional *to*, and complementiser *that* tends to be shorter and more reduced than pronominal *that*.

Phonetic differences that reflect grammatical differences can be perceptually significant. Shi et al. (1999) used the high-amplitude sucking paradigm to demonstrate that newborn infants (1 to 3 days old) were sensitive to the acoustic differences between a list of content words (e.g., *toys*, *chew*, *chair*, *find*) and a list of function words (*its*, *the*, *in*, *your*). The stimuli were recorded in isolation to prevent coarticulatory influences, and the speaker mimicked tokens from a recording of an English-speaking mother to her one-year-old child. Shi et al. suggest that infants' sensitivity to the acoustic differences between function and content words helps them to 'bootstrap' their way to acquiring syntactic categories.

Adult listeners can make use of these phonetic differences to facilitate perception. Baker (2008) demonstrated that mismatching phonetic detail between content and function words can disrupt perception. Sentences containing inappropriate phonetic detail were created by cross-splicing syllables from rhythmically-matched sentences containing phonemically-identical sequences in either function or content words (e.g., [ʃiːz] from *she's* spliced into the word *banshees*). These sentences produced fewer accurate responses in an intelligibility-in-noise task than control matched-splice sentences. Systematic phonetic detail had influenced how listeners interpreted the grammatical function of a phoneme sequence and so affected their interpretation of the syntactic structure of a sentence.

Evidence of associations between phonetic characteristics and grammatical category indicates that this type of association can be acquired during language acquisition, and is therefore a good candidate for demonstrating context-sensitive perceptual learning in adults.

Experiment 2a investigates whether listeners can learn to associate a duration difference with the grammatical categories of noun and verb. This is likely to be a learnable association because there is evidence to suggest that nouns and verbs have different phonological and phonetic properties which play a role in speech perception and even reading.

For example, the well-known distinction between first- and second-syllable stress for a small set of bisyllabic word pairs (*subject*, *survey*) extends to others; Sereno (1986) found that 92% of bisyllabic English nouns were stressed on the first syllable, while

85% of bisyllabic English verbs had stress on the second syllable (see also Kelly, 2004). Kelly and Bock (1988) showed that listeners are influenced by and make use of these prototypical stress patterns; subjects were more likely to pronounce bisyllabic pseudowords with first-syllable stress if they acted as nouns rather than verbs. Likewise, when subjects were asked to produce a sentence containing an auditorily-presented target bisyllabic pseudoword, they tended to use pseudowords as nouns if they had first-syllable stress, and as verbs if they had second-syllable stress (Kelly, 1988).

Sereno and Jongman (1995) investigated whether bisyllabic noun/verb homonyms that do not display the prototypical stress pattern may, nevertheless, exhibit systematic acoustic differences that reflect 'traces' of the prototypical stress pattern. Five subjects produced 16 grammatically ambiguous words (e.g., *answer* or *design*), read aloud as part of a list of nouns or a list of verbs. Noun and verb lists were controlled for stress placement in order to prevent bias due to list composition. Results showed small but consistent differences in duration and amplitude ratios between the first and second syllable, dependent on whether the token was intended as a noun or a verb. Listeners' sensitivity to these phonetic differences was not investigated.

Nouns and verbs also differ systematically in their number of syllables – as the number of the syllables in a word increases, it is more likely to be a noun than a verb. Cassidy and Kelly (1991) showed that subjects were sensitive to this tendency. The probability that subjects constructed a sentence that used a target pseudoword as a noun increased as the number of syllables in the pseudoword increased.

Vowel quality may also influence subjects' processing of nouns and verbs. In a corpus study of American English, Sereno and Jongman (1990) showed that, for high-frequency words, back vowels (57%) were more common than front vowels (43%) in the stressed syllable of nouns, and front vowels (62%) were more common than back vowels (38%) in the stressed syllable of verbs. (This skewed distribution did not apply to low-frequency words.) Reaction times in an auditory noun/verb categorisation task and an auditory lexical decision task reflected systematic processing differences between nouns and verbs depending on the quality of their stressed vowels: more typical words were processed faster, i.e. nouns with back vowels were processed

faster than nouns with front vowels, and, conversely, verbs with front vowels were processed faster than verbs with back vowels. This relationship was again only valid for high-frequency words. Sereno (1994) found a similar effect for a visual (reading) noun/verb categorisation task, although there was no interaction with frequency. Sereno and Jongman (1990, p. 402) conclude that subjects' expectations about the distribution of front and back vowels in relation to syntactic categories should be taken into account by researchers of auditory word recognition.

Durieux and Gillis (2001) assessed the usefulness of various phonological cues for predicting grammatical category. They used instance-based learning to train an artificial learning system to predict grammatical class based on exposure to the CELEXv2 database. No individual cue was a very strong predictor of grammatical class, but when a syllabified string of segments and the stress pattern were used in combination, almost 8 out of 10 cases were correctly classified as verb or noun. This suggests that sensitivity to statistical tendencies in the distribution of a combination phonological cues could be extremely effective in helping listeners to determine grammatical class.

Evidence that 'phonological typicality' affects the processing of nouns and verbs is presented by Farmer et al. (2006). They found that, although there was considerable overlap, monosyllabic nouns and verbs formed separate clusters in 'phonological feature space', i.e. when the distribution of phonological features was analysed, nouns were significantly closer to other nouns than would be expected by chance (and the same was found for verbs). Phonologically typical and atypical words were selected by calculating the distance of the word to all verbs and the distance of the word to all nouns. In a self-paced reading task, reaction times were slower for words that were phonologically less typical of the expected category, indicating on-line effects of phonological typicality on reading. In a similar task using words that could be interpreted as nouns or verbs, phonological typicality influenced readers' expectations of grammatical category, as measured by reaction times and comprehension accuracy. These experiments involved reading rather than speech perception, and so the influence of the phonological typicality of nouns and verbs on on-line spoken language comprehension has not yet been demonstrated. Nevertheless, the written word might be considered less closely related to phonological features than speech,

and so these results suggest that probabilistic differences in the phonological features of nouns and verbs could influence speech perception.

So, the verb-noun distinction seems to be probabilistically cued by syllable stress pattern, 'traces' of the stress pattern, syllable number, vowel quality and the distribution of phonological features. Experiment 2a investigates whether listeners can learn to associate a durational cue with the verb-noun distinction.

3.1.2 Experiment 2a: Methodological background

3.1.2.1 Selection of duration as the cue to be manipulated

Duration was manipulated because listeners are known to be extremely sensitive to durational cues, which can be associated with many kinds of distinction – segmental differences, prosodic structure, talker, speech rate, and possibly also emotional and pragmatic differences. In addition, listeners are known to adapt to durational cues such as VOT (Allen & Miller, 2004; Kraljic & Samuel, 2006).

Duration can also be manipulated with relative ease without creating speech that sounds unnatural. In the present experiment, duration was manipulated at a sub-phonemic level, i.e. without altering the phoneme percept. The use of a sub-phonemic distinction aimed to avoid the strategic responses that may have influenced results in Experiment 1, after which some subjects commented on the "funny 'I's" or that "the 'I's sounded like 'U's". In natural conversation, adaptation to a talker may involve an increase in that talker's intelligibility without (necessarily) a concomitant explicit awareness of how the talker makes certain distinctions.

3.1.2.2 Test task

Experiment 1 highlighted various problems with using a lexical decision task to assess perceptual learning (see Section 2.4.5). Experiment 2a aimed to avoid these difficulties by using a sentence completion task in which subjects had to interpret an ambiguous noun/verb (e.g., *the Indian <u>braves</u>…*) by typing their own ending to the sentence. For instance, *the Indian braves… fought the colonisers* indicates a noun interpretation of *braves*, while *the Indian braves… the weather* indicates a verb interpretation of *braves*.

Using this task should minimise learning during assessment, because either a noun or a verb interpretation of the keyword (e.g., *braves*) is possible, and neither is 'correct' or 'incorrect'. Therefore, in contrast to the lexical decision task, there is no top-down feedback from higher levels of linguistic structure (e.g., word knowledge) that can guide learning during the test task.

Unlike a lexical decision task, the sentence completion task does not demand an explicit meta-linguistic judgement, but instead involves interpreting the meaning of the speech stimulus. This means that it should resemble more closely the listeners' task in understanding speech in everyday conversation.

3.2 EXPERIMENT 2A: METHOD

3.2.1 Subjects

56 subjects (24 male, 32 female) participated in the experiment. They were non-phonetically-trained native speakers of British English, monolingual (until at least age 11), aged between 18 and 30, and had no known speech or hearing problems. Nearly all were students at the University of Cambridge. They were unfamiliar with the voice of the talker, and were paid £6 for their participation.

3.2.2 Materials

3.2.2.1 Training materials: Story

Training material consisted of an original story read by a 60-year-old female phonetician in a Standard Southern British English (SSBE) accent. The speaker was recorded in a sound-treated room using a Marantz PMD670 solid-state recorder at a sampling rate of 22050 Hz using a Sennheiser MKH 40 P48 high-frequency cardioid condenser microphone. The story lasted approximately 12 minutes, and the transcript is in Appendix E.

The story contained 326 nouns and 326 main verbs. Constructions which could be considered ambiguous in terms of their inclusion of a noun or a verb were avoided. This excluded any gerunds (e.g., *Dancing is fun*), and ambiguous past participles which could be interpreted as adjectives e.g., *The plane was delayed*. All pronouns, auxiliary verbs and modal verbs were in unstressed positions in the sentence, and their

durations were not manipulated. Compound nouns were avoided because of the difficulty determining whether to manipulate the duration of one word or both words.[5]

Two versions of the story were created from the original recording. For the **NSVL** version, stressed vowels in nouns were shortened by 25% and stressed vowels in verbs were lengthened by 33%. For the **NLVS** version, stressed vowels in nouns were lengthened by 33% and stressed vowels in verbs were shortened by 25%. Informal auditory assessment by the author and the speaker indicated that altering the duration of the stressed vowel (i.e. the periodic syllable nucleus, excluding adjacent consonant transitions) gave a more natural-sounding accent than lengthening or shortening the entire verb/noun or the entire stressed syllable. This is consistent with evidence that the compression of consonants is more disruptive to perception than the compression of vowels (Gordon-Salant & Fitzgibbons, 2001), and with the greater effect of rate of speech on vowel duration than consonant duration (Pickett, 1999). Shortening vowels to less than 75% of their original length led to an unnatural percept, but vowels could be lengthened to 133% without sounding unnatural. (Lengthening/shortening factors are only approximate, because the precise amount of shortening/lengthening is dependent on the period of the vowel.) The duration of unstressed vowels was not manipulated because unstressed vowels were often highly reduced in the training story and were therefore difficult to identify and label, and because the duration manipulation gave a relatively natural-sounding percept when restricted to stressed vowels. For the criteria used to label stressed vowels in nouns and verbs, see Appendix F.

Durations were manipulated using Time-Domain Pitch-Synchronous Overlap-And-Add (TD-PSOLA) resynthesis in Praat (Boersma & Weenink, 2007). TD-PSOLA aims to manipulate the prosody of speech while retaining a high degree of naturalness (Moulines & Charpentier, 1990). The quality of the resynthesis was checked auditorily and by inspecting the spectrograms for discontinuities. Where creak in the original recording caused discontinuities, the affected sentences were re-recorded until an acceptable resynthesis was achieved. The speaker aimed to match rate, pitch and prosody to the original recording.

[5] One compound noun *Isabel Davis* was included near the beginning of training because it was necessary for the flow of the story. The stressed vowels in both words were manipulated.

Comprehension questions (see Appendix G) were included to ensure that subjects paid attention to the story, and to serve as a distracter from the true purpose of the experiment.

3.2.2.2 Test materials: Sentence completion

Perceptual learning was assessed using a sentence completion task. The same talker was used in training and testing. There were 20 experimental sentence beginnings and 80 filler sentence beginnings. Experimental sentence beginnings contained a novel ambiguous noun/verb keyword. They were constructed such that the second word could be interpreted as an adjective or a noun, and the third word (the keyword) could be interpreted as a noun or a verb e.g., *the Indian* (Adj/N) *braves* (N/V) …

All sentences beginnings were recorded as part of a complete sentence, and then truncated at the end of the keyword. The influence of coarticulation on subjects' responses to the ambiguous sentence beginnings was minimised by immediately following the noun/verb with a schwa, and by avoiding /r/ (which can have long-domain resonances) for two syllables after the keyword. A 'neutral' (H* L) prosody was used such that both noun and verb interpretations of experimental sentences were possible. To help ensure this neutral prosody, a noun version and a verb version of each experimental sentence was recorded. These sentences (listed in Appendix H) were matched for nuclear stress position, number of syllables, and, as far as possible, for segment quality. Any sentence beginnings with audible differences in pitch or duration, or easily discernible spectrographic differences in duration or F0 traces were re-recorded. The final stimuli were selected such that half were originally recorded with a verb interpretation and half with a noun interpretation.

Sentence beginnings were piloted to ensure that they were sufficiently ambiguous. Only sentence beginnings achieving between 25% and 75% noun interpretations (and hence also 25–75% verb interpretations) were included. These are listed in Table 2 alongside the pilot results.

Set A	Number of valid pilot responses	Noun responses (%)	Set B	Number of valid pilot responses	Noun responses (%)
the Indian **braves**	16	56	the government **houses**	8	50
the college **grants**	16	63	the factory **fires**	16	69
the French **cook**	16	69	the animal **calls**	8	63
the police **report**	16	69	the army **supplies**	16	56
the primrose **flowers**	16	69	the leopard **spots**	16	75
the official **lies**	16	44	the guest **reviews**	16	63
the major **fights**	15	60	the British **show**	8	63
the Russian **pines**	16	75	the chief **demands**	8	50
the community **notices**	8	50	the dwarf **plants**	16	69
the graduate **plays**	7	57	the terrorist **states**	7	57
Mean noun responses (%)		61	**Mean noun responses (%)**		61

Table 2. Experimental sentence beginnings used in the sentence completion task, alongside noun responses (%) from pilot experiments. Ambiguous noun/verb keywords are in bold. The two halves of the table represent sets used for counterbalancing stimuli across groups (see Section 3.2.3).

Two versions of each experimental sentence were created such that the stressed vowel in the ambiguous noun/verb was either lengthened by 33% or shortened by 25%. Henceforth, lengthened keywords will be referred to as **Long** and shortened keywords as **Short**. The duration of the stressed vowel in the second word was left unchanged because it can be interpreted as a noun or an adjective, e.g. *the Indian*(N) *braves*(V) or *the Indian* (Adj) *braves* (N), and therefore it was impossible to choose a duration (Long or Short) which was consistent with both noun and verb interpretations of the keyword. It was hoped that the duration of the second word may not have much influence on its interpretation as an adjective or a noun because the first word was always the short function word *The*. Subjects may not have formed a strong impression of the speech rate by the second word, and so both noun and verb interpretations of the keyword should remain plausible.

The 80 filler sentence beginnings contained 64 nouns and 64 verbs between them. A NLVS and NSVL version of each sentence was created. Appendix J lists the fillers.

Due to technical problems, the recording equipment differed slightly from that used for the training story. Test materials were recorded using Media Recorder software on a Silicon Graphics computer. There was no discernible difference in recording quality.

3.2.3 Procedure

Testing conditions, equipment and presentation software were identical to Experiment 1. Subjects were presented with written instructions which were clarified by the experimenter as necessary. There was no practice session.

Subjects were assigned randomly to one of two Training groups. 28 subjects heard only the NLVS training story and 28 subjects heard only the NSVL training story. Training and test tasks were interleaved 10 times throughout the experiment, in order to minimise the chance that a period of exposure to truncated sentences might eliminate any learning effect from the continuous speech in the story.

The instructions requested subjects to listen carefully to the training story, bearing in mind that they would be asked about it later in the experiment. The first section of story lasted 5 minutes, in order to allow a reasonable amount of time for learning to occur before the first test task. The other nine sections lasted an average of 52 seconds.

In the test task, subjects were requested to listen to the beginning of the sentence and then type out their own ending to the sentence. They were asked to respond as quickly as possible with the first thing that 'springs to mind', and were told not to worry about writing sentences with a sensible meaning so long as they made grammatical sense.

The experimental sentence beginnings were split into two sets, Set A and Set B. Set A and Set B were matched as far as possible for ambiguity (based on the pilot results) and phonological vowel length (see Table 2). Within each Training group (NSVL or NLVS), half of subjects heard Set A words lengthened and Set B words shortened, while the other half heard Set A words shortened and Set B words lengthened. Subjects heard either NSVL or NLVS versions of the fillers, to match the accent heard during training.

Each section of the sentence completion task comprised 10 sentence beginnings (2 experimental and 8 fillers). Sentence beginnings were randomised with the constraint

that two experimental sentences occurred within each section of the sentence completion task, and these could not be consecutive stimuli. The order of stimuli was counterbalanced within each condition (e.g., NSVL-SetALong) to create 8 orders in total.

A comprehension question was included at the end of alternate test tasks. The comprehension questions required a short typed response.

The experiment took between 35 and 55 minutes to complete (times varied depending on the length of subjects' responses and their typing speed).

3.3 EXPERIMENT 2A: HYPOTHESIS

It was predicted that the accent heard during training would influence listeners' perceptual decisions. Therefore, an interaction between Training group and Duration was predicted: NLVS subjects were expected to interpret more Long words as nouns and more Short words as verbs than NSVL subjects.

The likely rate of learning was uncertain, and all measurable learning could have occurred before the first task, resulting in no effect of Task number. However, if learning occurred over the course of the experiment, responses were expected to become more consistent with the accent heard during training as Task number increased.

3.4 EXPERIMENT 2A: RESULTS

3.4.1 Data pre-processing

Responses to the ambiguous sentence beginnings were coded as verb interpretations or noun interpretations. 3% of responses were excluded from the analysis because they were missing, demonstrated misunderstanding of the sentence beginning, or were ambiguous. 89% of comprehension questions were answered correctly, with subjects ranging from 40–100% correct.

3.4.2 Analysis of verb and noun responses

The data were analysed using generalised linear mixed effects modelling, with fixed predictors of Training group (NSVL, NLVS), Duration (Short, Long), Task number (1 to 10) and all interactions, and random factors of Item and Subject. The most

parsimonious model included a significant interaction between Duration and Training group ($\chi^2 = 6.2$, $df = 1$, $p < 0.02$), and between Task number and Training group ($\chi^2 = 4.7$, $df = 1$, $p < 0.04$).

Pairwise comparisons (see Figure 4) showed that NLVS subjects interpreted significantly more Long words than Short words as nouns (69% vs. 53%, $z = 4.7$, $p < 0.001$). In addition, as predicted, significantly more Long words were interpreted as nouns by NLVS subjects than by NSVL subjects (69% vs. 59%, $z = 2.0$, $p < 0.04$). In contrast, the effect of Duration did not emerge as significant for NSVL subjects (54% for Short words, 59% for Long words, $z = 1.4$, $p = 0.2$), and there was no significant difference in responses between NLVS and NSVL subjects for Short words, although there was a non-significant trend in the predicted direction, with NSVL subjects interpreting a slightly higher percentage of Short words as nouns than NLVS subjects (54% vs. 53%, $z = 0.3$, $p = 0.8$).

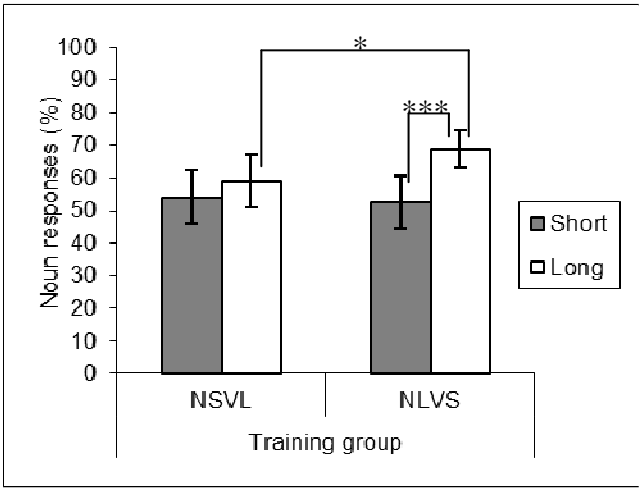


Figure 4. Percentage of noun responses in the sentence completion task, as a function of Duration and Training group. Error bars show the 95% confidence interval of the mean. Significant pairwise comparisons are indicated as follows: * $p < 0.05$; *** $p < 0.001$.

There was a significant interaction between Task number and Training group (see Figure 5); NLVS subjects gave fewer noun responses as Task number increased ($z = 3.8$, $p < 0.001$), while there was no effect for NSVL subjects ($z = 0.8$, $p = 0.4$).

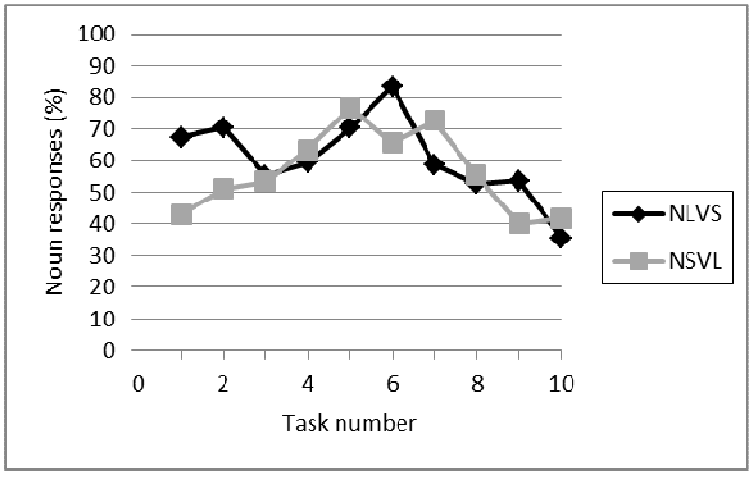


Figure 5. Percentage of noun responses in the sentence completion task, as a function of Task number and Training group.

3.5 EXPERIMENT 2A: DISCUSSION

In this experiment, listeners adapted to a sub-phonemic durational cue that was related to grammatical class, and used their newly-acquired knowledge to guide their interpretation of syntactic structure. NLVS subjects were more likely than NSVL subjects to interpret Long words as nouns in the sentence completion task. The lack of any significant effect of Training group on the interpretation of Short words was unexpected, and is discussed further in Section 3.11.1 in the light of the results of Experiment 2b. The decreasing number of noun responses for NLVS subjects as the experiment progressed was not predicted, and may be task-related, as is discussed in Section 3.10.2.

Although the results for Long words demonstrate that the duration of nouns and verbs heard during training can affect subjects' interpretation of an ambiguous noun/verb, it is not possible to conclude that subjects learned to associate a durational cue directly with the grammatical classes, noun and verb. The grammatical category of a word and its position in prosodic structure are somewhat confounded, and so this study cannot distinguish learning that relates to grammatical category from learning that relates to prosodic phrase position.

In English, nouns are always at the end of a syntactic noun phrase, while verbs are sometimes at the end of a syntactic phrase (e.g., [*The dog*$_{NP}$] [*sleeps*$_{VP}$]) and sometimes not (e.g., [*The dog*$_{NP}$] [*bites* [*the cat*$_{NP}$] $_{VP}$]). This close relationship between syntactic and prosodic structure has the consequence that nouns are more often at the end of prosodic phrases than verbs. Listeners could therefore have learned either an association between the durational cue and nouns and verbs, or an association between the durational cue and prosodic phrase position.

The assessment task in Experiment 2a is unable to distinguish these two possibilities because it also confounds grammatical class and syntactic phrase position. The structure of the ambiguous sentence beginnings is such that, where the keyword is interpreted as a noun, it is always at the end of a syntactic phrase. In contrast, when it is interpreted as a verb, the keyword is not at the end of a syntactic phrase (see Figure 6 for an example).

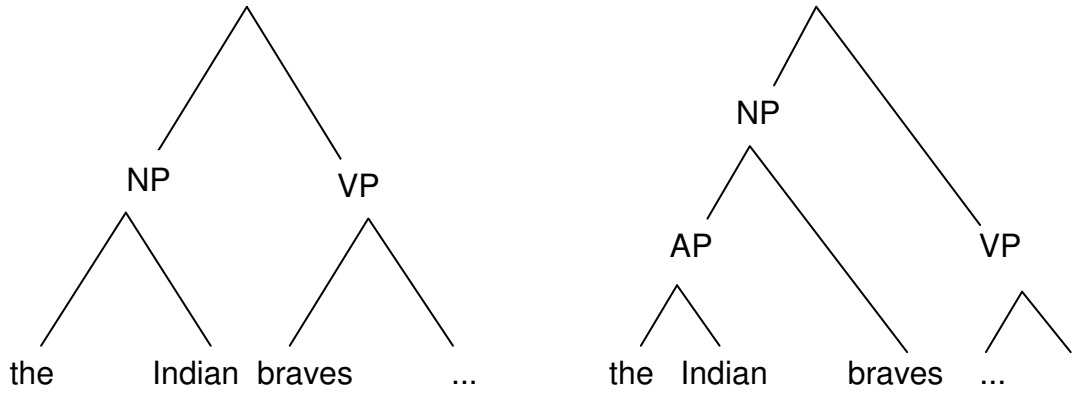


Figure 6. Syntactic trees for verb (left) and noun (right) interpretations of *the Indian braves ...*

The sentence beginnings were intentionally produced such that their interpretation was ambiguous (see Section 3.2.2.2). That is, there was no clear phonological phrase boundary signalling where the noun phrase ended. (Following Shattuck-Hufnagel and Turk (1996), the term 'phonological phrase' (PP) will be used to refer to prosodic phrases within the intonational phrase.) Duration is known to affect subjects' perception of intonational phrase boundaries (e.g., Nagel et al., 1996), and prosodic cues have also been shown to affect the interpretation of phonological phrase boundaries (Christophe et al., 2004; Millotte et al., 2007, 2008). Millotte et al. demonstrated that French listeners used prosodic cues (e.g., pitch contour and duration) to help interpret syntactically ambiguous verb/adjective structures such as [*le petit chien*$_{PP}$] [*mord...*] ('the little dog bites...') versus [*le petit chien mort*$_{PP}$] ('the dead little dog...'). Their studies used French and an adjective/verb category ambiguity rather than a verb/noun one, but nevertheless it seems likely that that English listeners can use prosodic cues such as duration to disambiguate structures of the type encountered in the sentence completion task in Experiment 2a. This suggests that if training affected subjects' knowledge of how the talker signals prosodic phrase boundaries using duration, it may have influenced their perception of the presence or absence of a phonological phrase boundary immediately after the noun/verb keyword.

Experiment 2b aimed to determine whether listeners learned to associate the durational cue with prosodic phrase position during training, as described above, or whether learning was directly related to the grammatical categories of noun and verb.

3.6 EXPERIMENT 2B: METHODOLOGICAL BACKGROUND

It was not possible to construct a training story that eliminated the confound between grammatical category and prosodic position, therefore Experiment 2b aimed to clarify what was learned from training, rather than attempting to induce a definitive association between duration and grammatical class during training. Experiment 2b replicated Experiment 2a, but used an assessment task in which ambiguous sentence beginnings were structured such that the critical noun/verb was always in phrase-final position.

The stimuli in the sentence completion task in this experiment are based on ambiguous noun/verb stimuli constructed by Cooper and Paccia-Cooper (1980). They investigated whether the observation that nouns are longer than verbs could be explained purely as a phrase-final lengthening effect. Phonetically-controlled sentences in which homonym nouns and verbs always occurred at the end of an intonational phrase were recorded. Results showed no systematic durational differences between nouns and verbs when prosodic position was controlled. A sample sentence pair is given below:

Noun: *John will find Eve a <u>coach</u> if she ever decides to sing professionally*

Verb: *John will help Eva <u>coach</u> if she ever decides to start a basketball team*

The sentence completion task in Experiment 2b used a variation on this, such that the same sentence beginning was compatible with either verb or noun interpretations. The experimental stimuli contained the verb *make* and names that can have or lack a final schwa, depending on whether the critical word is interpreted as a noun or a verb, e.g. *I made [Peter/Pete a] <u>plan</u> although...* For these sentence beginnings, the ambiguous keyword is prosodically phrase-final whether interpreted as a noun or a verb.

3.7 EXPERIMENT 2B: METHOD

3.7.1 Subjects

69 subjects (18 male, 51 female), aged 18–30, participated. Criteria for participation were identical to Experiment 2a. Five subjects (1 male, 4 female) were excluded (see

Section 3.9.1), leaving 64 subjects (32 in each Training group) in the statistical analysis.

3.7.2 Materials

3.7.2.1 Training materials: Story

The training story and comprehension questions were identical in Experiments 2a and 2b.

3.7.2.2 Test materials: Sentence completion

The 20 experimental sentence beginnings are listed in Table 3. Each contains one novel ambiguous noun/verb keyword in phrase-final position.

Set A	**Set B**
He'll make [Norma/Norm a] **bet** if	She made [Joanna/Joanne a] **dress** so that
They made [Suzannah/Suzanne a] **drink** whenever	They made [Philippa/Philip a] **drum** because
We made [Elena/Elaine a] **pack** because	I made [Peter/Pete a] **plan** although
They made [Diana/Diane a] **print** in case	I made [Christina/Christine a] **snack** whenever
I'll make [Paula/Paul a] **deal** if	He'll make [Eva/Eve a] **bowl** if
He made [Maria/Marie a] **sign** because	They made [Gina/Jean a] **file** since
He made [Daniella/Danielle a] **train** because	I made [Louisa/Louise a] **spy** even though
They'll make [Helena/Helen a] **budget** so	He made [Tricia/Trish a] **copy** since
We made [Georgia/George a] **model** so	They made [Jonah/Joan a] **promise** so
They'll make [Rosa/Rose a] **recruit** until	They made [Carla/Carl a] **report** because

Table 3. Experimental sentence beginnings used in the sentence completion task. Ambiguous noun/verb keywords are in bold. The two halves of the table represent sets used for counterbalancing stimuli across groups (see Section 3.7.3).

In order to neutralise any phonetic cues to the presence or absence of a word boundary between the name and the schwa (e.g., *Peter* vs. *Pete a*), a noun version and a verb version of each sentence was recorded (see Appendix I), then truncated, and an intermediate version was created by merging the noun and verb versions using STRAIGHT (Kawahara, 2006). STRAIGHT is a system for analysing, modifying and resynthesising speech, and can be used to blend two original speech tokens in differing proportions to create natural-sounding intermediate stimuli.

The experimental stimuli were piloted to ensure that they were ambiguous. An initial pilot used stimuli which were weighted (in STRAIGHT) to be half-way between the

original noun version and the original verb version. 7 subjects took part in a sentence completion task using these stimuli, but this failed to identify enough ambiguous stimuli.

In a second pilot experiment, nine versions of each sentence beginning were created using STRAIGHT; one version based on the original noun version, one based on the original verb version, and seven intermediate steps. These were randomised with constraint that same sentence fragment did not occur consecutively. 8 subjects took part in a 2AFC categorisation task. Sentence beginnings for which no step on the noun–verb continuum reached above 40% noun responses or below 60% noun responses were excluded. For the remaining sentence beginnings, the stimulus which achieved closest to 50% noun responses was selected (where there were two equal candidates, the stimulus that was physically closest to half-way along the verb–noun continuum was selected). Stressed vowels in the ambiguous verb-nouns were either lengthened by 33% or shortened by 25% using Praat TD-PSOLA resynthesis.

There were 80 filler sentences. Most of these were the same as in Experiment 2a. However, the experimental sentence beginnings in Experiment 2b were longer (i.e. contained more words) than those in 2a and all contained proper names, so some of the Experiment 2a fillers were replaced in order to better 'disguise' the experimental stimuli in Experiment 2b. Experiment 2b fillers are in Appendix J, and contained 91 main verbs and 91 nouns in total. An NSVL and an NLVS version of each filler sentence was created using Praat TD-PSOLA resynthesis.

3.7.3 Procedure

35 subjects were randomly assigned to the NLVS Training group, and 34 to the NSVL Training group. The procedure was the same as in Experiment 2a, with the exception that subjects were requested to type the sentence beginning followed by their own sentence completion. This differs slightly from Experiment 2a, in which subjects had to type only their sentence completions. This change was necessary because the same sentence endings can be equally appropriate for both verb and noun interpretations of some sentence beginnings, e.g. *They made* [*Christina/Christine a*] *snack whenever... she was hungry*.

The stimuli were split into Set A and Set B, balanced as far as possible for syllable number and phonological vowel length (see Table 3). As for Experiment 2a, half of the subjects within each Training group heard Set A words lengthened and Set B words shortened, while the other half heard Set A words shortened and Set B words lengthened. Subjects heard either NSVL or NLVS versions of the fillers, to match the accent heard during training. The randomisation for Experiments 2a and 2b was identical, i.e. experimental stimuli from Experiment 2a were directly replaced by those from Experiment 2b, and certain fillers were replaced as indicated in Appendix J.

3.8 EXPERIMENT 2B: HYPOTHESIS

Experiment 2b sought to clarify what was learned by subjects in Experiment 2a. If the perceptual learning effect was due to learning an association between a duration difference and grammatical status, NLVS subjects should interpret more Long words as nouns and more Short words as verbs than NSVL subjects. If the learning effect was related to phrase position, there should be no interaction between Training group and Duration.

3.9 EXPERIMENT 2B: RESULTS

3.9.1 Data pre-processing

5 subjects were excluded for failing to follow instructions correctly. 32 NSVL and 32 NLVS subjects were included in the statistical analysis.

Responses to the ambiguous sentence beginnings were coded as verb interpretations or noun interpretations. The 5% of responses that demonstrated misunderstanding of the sentence beginning were excluded. 89% of comprehension questions were answered correctly, with subjects ranging from 40–100% correct.

3.9.2 Analysis of verb and noun responses

The full mixed model had fixed factors of Training group (NLVS, NSVL), Duration (Long, Short), Task number (1 to 10) and all interactions, and random factors of Subject and Item. Only Task number was a significant predictor in the most parsimonious model ($\chi^2 = 31.9$, $df = 1$, $p < 0.001$): a higher proportion of noun responses were given as the experiment progressed. Performance in each subject

group was virtually identical: both NSVL subjects and NLVS subjects interpreted 51% of Long stimuli as nouns, and 49% of Short stimuli as nouns.

3.10 EXPERIMENT 2B: DISCUSSION

The lack of an interaction between Training group and Duration in this experiment suggests that the significant effect in Experiment 2a was due to subjects learning to associate a duration difference with prosodic phrase position rather than directly to grammatical category. When phrase position was controlled in the test task in Experiment 2b, the effect disappeared. It seems unlikely that other differences in the experimental stimuli or procedure between Experiments 2a and 2b can account for these results, as is discussed in Section 3.10.1.

3.10.1 Differences between Experiment 2a and Experiment 2b

3.10.1.1 Procedural difference

There was just one difference in the subjects' task between the experiments: subjects in Experiment 2b had to type the sentence beginning, as well as their own sentence completion. The additional time available in Experiment 2b before subjects started to type their own ending may have influenced the precise form of that ending, but it seems unlikely to have influenced subjects' interpretation of the sentence beginning. Additional memory load could have led 2b subjects to focus more on the sentence beginnings, but this seems unlikely to have made Experiment 2b *less* sensitive than Experiment 2a.

3.10.1.2 Differences in experimental sentence beginnings

In contrast to Experiment 2a, the noun/verb keyword in sentence beginnings in Experiment 2b was always at the end of a prosodic phrase. However, there were inevitably other differences between the experimental stimuli. This section briefly discusses why these differences are unlikely to account for the null effect of Training group in Experiment 2b.

Stimulus ambiguity

The frequency and transitional probabilities of the ambiguous words are likely to have affected subjects' responses in Experiment 2a and Experiment 2b. For example, a noun response may be probable if a subject has heard *French cook* more frequently

with a noun than a verb interpretation of *cook*, or if *Joanne a* has been heard more frequently than *Joanna*. A strong bias towards either noun or verb responses for individual items could have limited the potential for shifting responses between noun and verb interpretations, and could therefore potentially account for the lack of an effect in Experiment 2b. This possibility was minimised by including only sentence beginnings that achieved 25–75% noun responses in the pilot experiment. Table 4 indicates that, although the percentage of noun responses in Experiment 2b fell outside the 25–75% window for a few items, the majority of sentence beginnings had a lot of both verb and noun responses. Furthermore, the significant effect of Task number indicates that the interpretation of the stimuli was malleable.

Sentence beginnings	**Noun responses (%)**	**Sentence beginnings**	**Noun responses (%)**
They made [Jonah/Joan a] **promise** so	86	He made [Daniella/Danielle a] **train** because	52
We made [Georgia/George a] **model** so	80	She made [Joanna/Joanne a] **dress** so that	51
They'll make [Helena/Helen a] **budget** so	77	We made [Elena/Elaine a] **pack** because	40
They made [Gina/Jean a] **file** since	76	They made [Diana/Diane a] **print** in case	37
They'll make [Rosa/Rose a] **recruit** until	68	I made [Louisa/Louise a] **spy** even though	36
I'll make [Paula/Paul a] **deal** if	67	He'll make [Norma/Norm a] **bet** if	34
I made [Christina/Christine a] **snack** whenever	62	I made [Peter/Pete a] **plan** although	28
He'll make [Eva/Eve a] **bowl** if	58	They made [Carla/Carl a] **report** because	27
They made [Philippa/Philip a] **drum** because	54	They made [Suzannah/Suzanne a] **drink** whenever	24
He made [Tricia/Trish a] **copy** since	53	He made [Maria/Marie a] **sign** because	9

Table 4. Mean percentage of noun responses for each sentence beginning. Ambiguous noun/verbs are in bold.

Stimulus similarity

The experimental stimuli in Experiment 2b all used the verb 'make', which could have alerted subjects to the nature of the experiment and encouraged more strategic responses, i.e. adopting a noun response mode or a verb response mode. However, a comparison of subject scores in each experiment (see Figures 7 and 8) shows that there were a substantial number of subjects with very mixed responses in both experiments.

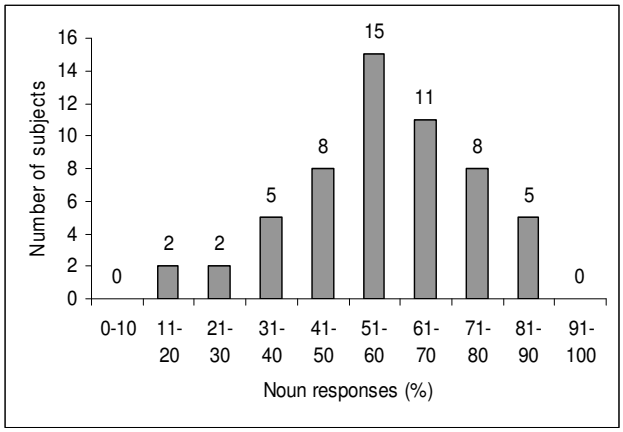


Figure 7. Number of subjects within each 10% 'bin' of noun responses in the sentence completion task in Experiment 2a.

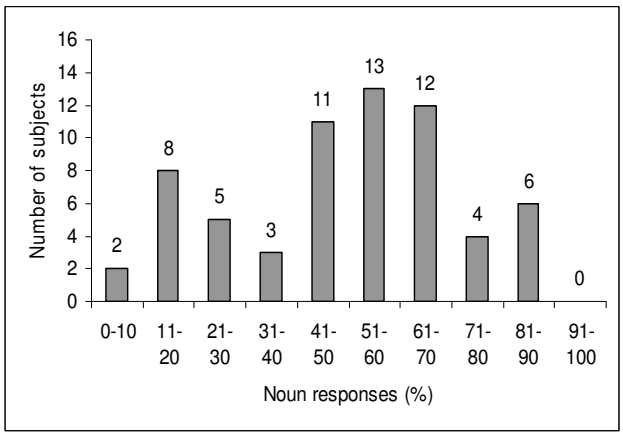


Figure 8. Number of subjects within each 10% 'bin' of noun responses in the sentence completion task in Experiment 2b.

Resynthesis

The stimuli in Experiment 2b were resynthesised twice, once by STRAIGHT and once using Praat (PSOLA), whereas the stimuli in Experiment 2a were resynthesised just once. However, 'double resynthesis' did not (in the experimenter's judgement) cause a noticeable difference in the quality of the synthesis.

Sentence length

Likewise, sentence length prior to the critical word should not have caused the observed difference. Experiment 2a had an average of 3.15 syllables prior to the keyword, while Experiment 2b had 4.55 syllables. If the number of prior syllables affected subjects' ability to judge the rate of speech, the assessment of keyword duration should have been *more* accurate for Experiment 2b.

Lack of duration manipulation prior to keyword

In both experiments, duration was only manipulated on the critical ambiguous noun/verb in the experimental sentence beginnings, for the reasons discussed in Section 3.2.2.2. This meant that where the keyword was interpreted as a verb in Experiment 2a, the word immediately prior to the keyword (e.g. *Indian* from *The Indian braves*…) was a noun, but did not have the appropriate duration (Long or Short) consistent with the accent heard in training. In Experiment 2b, under both noun and verb interpretations of the keyword, the verb ('make') and a proper noun (the name) had inappropriate durations as compared to those heard in training. This lack of consistency could have caused subjects to dissociate test from training and limited the detection of any effect of Training group. This dissociation may have been more likely in Experiment 2b because the experimental stimuli always contained a verb and a proper noun before the keyword.

However, the lack of duration manipulation on words before the keyword did not prevent subjects applying their newly-acquired knowledge in Experiment 2a, so it seems unlikely that slightly less accent consistency in Experiment 2b would have *totally eliminated* the application of this knowledge (as the results of 2b indicate). In contrast, the elimination of a Training effect is expected if subjects learned about how the talker signals prosodic boundaries, so the latter explanation is more probable.

3.10.2 Effect of Task number

The effect of Task number in Experiments 2a and 2b may be a task-related change in behaviour. In Experiment 2b subjects gave more noun responses as the experiment progressed, whereas in Experiment 2a NLVS subjects gave more verb responses as the experiment progressed. These results may be explained by changing 'activation levels' for noun and verb interpretations as subjects' experience of the experimental stimuli increased. If subjects became aware of certain noun/verb ambiguities as the experiment progressed, they may have been more likely to consider (either consciously or unconsciously) both noun *and* verb interpretations of later experimental stimuli. This would have reduced the bias of particular stimuli to generate either a noun or verb interpretation. Stimuli that produced a bias towards noun responses at the beginning of the experiment would show a less strong bias later in the experiment, as the results of NLVS subjects in Experiment 2a suggest. In

contrast, stimuli that produced a bias towards verb responses at the beginning would have more noun responses later in the experiment, as shown by the results of Experiment 2b. Subjects' awareness of some noun/verb ambiguities in informal post-experiment questioning suggests this explanation is plausible.

3.11 GENERAL DISCUSSION OF EXPERIMENTS 2A AND 2B

These experiments did not show perceptual learning relating to grammatical category, but instead demonstrated context-sensitive perceptual learning that related to prosodic phrase position. Learning to associate a sub-phonemic durational cue with the way a talker signals phrase boundaries shows that perceptual learning need not involve adjustments to phoneme boundaries or learning to associate a phonemic or featural realisation with a particular talker. A model of speech perception must account for listeners' ability to adapt to sub-phonemic information that relates prosodic structure.

3.11.1 Why was there no learning effect for Short words in Experiment 2a?

Having established that listeners learned about prosodic phrase boundaries rather than grammatical class, it is possible that the lack of any effect of Training group on Short words in Experiment 2a can be explained by an asymmetric effect of phrase-final lengthening. That is, if phrase-final lengthening is there, it indicates a boundary; if it is not, there might or might not be a boundary so listeners will perform at chance or based on prior expectations. This is consistent with the proposal that the presence of a prosodic boundary has a larger effect than the absence of a prosodic boundary (Gerken, 1996; Schepman & Rodway, 2000).

3.11.2 A role for prior associations in constraining perceptual learning

Experiments 2a and 2b suggest that listeners can more easily learn about a durational cue that reflects prosodic structure than a durational cue that is associated directly with grammatical category. In the training story, the relationship between grammatical category and the duration difference was categorical (i.e. 100% of main verbs were lengthened and nouns were shortened, or vice versa), whereas the durational difference and prosodic phrase position were merely correlated. Nevertheless, it was phrase position that became associated with the duration difference. This may be due to the influence of prior knowledge, as discussed below.

It is well known that prosody is critical to perception (see also Section 1.3.3). Phrase-final lengthening has been identified in production experiments since the 1970s (e.g., Martin, 1970; Klatt, 1975), and there is mounting evidence for its role in the resolution of syntactic ambiguities (e.g., Lehiste et al., 1976; Price et al., 1991; Speer et al., 1996; Nagel et al., 1996). Duration is therefore an established cue to prosodic structure, and the results of Experiments 2a and 2b suggest that this prior association had a strong influence on what was learned.

In contrast, although there is evidence for systematic phonetic differences between nouns and verbs, these cues are probabilistic, relatively subtle, and seem to be more effective as part of a 'constellation' of cues (as discussed in Section 3.1.1).

Therefore, it seems probable that where there are already strong associations between phonetic cues (or types of cue) and particular structures, listeners may be especially sensitive to changes in these regularities. This will be discussed further in Chapter 5 in the context of proposed properties for a model of perceptual learning.

3.11.3 What was learned in Experiment 2a?

The results of Experiment 2b indicate that the perceptual learning effect in Experiment 2a was due to subjects learning an association between the durational cue and prosodic phrase position. However, it is not possible to determine whether NLVS subjects learned that this talker exaggerates phrase-final lengthening, and/or whether NSVL subjects learned that the phrase-final lengthening cue is reduced or unreliable for this talker.

Either possibility seems plausible. It may be easy to learn that a talker exaggerates a particular cue with which the listener is familiar (e.g., phrase-final lengthening) because this is compatible with prior experience. On the other hand, a talker who goes against the norm may attract more attention to their atypical realisations, making perceptual learning likely. Impressionistically, the NSVL training passage sounded slightly more unusual than the NLVS condition.

Previous studies of perceptual learning indicate that both characteristics that are consistent with prior knowledge, and those that violate expectations are learnable. A Control condition with no duration manipulations during training would have helped to resolve the question of what was learned in Experiments 2a.

3.12 CONCLUSION

Experiments 2a and 2b have demonstrated context-sensitive perceptual learning relating to prosodic structure that can influence the interpretation of grammatical structure and syntactic class. This provides important support for the hypothesis that listeners can learn novel phonetic associations with higher levels of linguistic structure. In addition, perceptual learning of an association between a durational cue and prosodic phrase position rather than grammatical class suggests that there is a strong role for prior knowledge in what is learned or learnable.

It remains to be established, however, whether listeners are able to adapt to phonetic detail that is conditioned by grammatical class. Experiment 3 is a further investigation of context-sensitive perceptual learning conditioned by an aspect of grammar, although it investigates morphological rather than syntactic distinctions.

CHAPTER 4

EXPERIMENT 3: PERCEPTUAL LEARNING OF PHONETIC DETAIL THAT INDICATES MORPHOLOGICAL STRUCTURE OF WORDS

4.1 BACKGROUND

As discussed in Section 3.11.2, the existence of a prior association between phonetic detail and a particular structure may make it easier to learn new phonetic associations with that structure. With this in mind, Experiment 3 explored grammatically-sensitive perceptual learning further by focussing on a structural level of representation for which there is a relatively well-established association with phonetic detail. Specifically, it investigated perceptual learning of phonetic detail that indicates the morphological structure of words.

4.1.1 Phonetic detail and morphology

Several studies provide evidence that phonetic detail can reflect morphological structure. For example, in a small-scale study, Walsh and Parker (1983) found longer duration of morphemic /s/ compared with non-morphemic /s/ in sentences such as *Wrecks bothered him* vs. *Rex bothered him* (although word frequency was not controlled between the monomorphemic and bimorphemic word-pairs). More recently, Cho (2001) demonstrated that morpheme boundaries can affect intergestural timing in Korean. Using electromagnetic articulography, Cho compared the realisation of /pi/ in monomorphemic words such as /sapi/ ('private expense') and when it crosses a morpheme boundary, for example, /sap-i/ ('shovel' + Nom.). The duration from the midpoint of the labial closure to the end of the vowel was more variable when the sequence crossed a morpheme boundary. A second experiment supported this result: using electropalatography to investigate palatalisation in /ti/ and /ni/ sequences, Cho found more variability in the timing of articulatory gestures that crossed a morpheme boundary.

Hay (2003) investigated the relationship between the phonetic realisation of a /t/ that precedes a suffix, and the 'decomposability' of that suffix. Decomposability was

varied by manipulating the relative frequency of the bare stem as compared to the suffixed form. For example, *swiftly* is more frequent than *swift*, and is therefore less decomposable than *softly*, which is much less frequent that *soft*. Speakers of American English showed more /t/ reduction or deletion when a word was less decomposable (e.g., /t/ was shorter in *swiftly* than *softly*). Thus, there was a close relationship between the decomposability of the morpheme boundary and the phonetic realisation.

Baker (2008) provides further evidence that morphological structure can be reflected in phonetic detail. Baker compared the acoustic characteristics of prefixes and 'pseudo-prefixes', which start with the same phonemic sequence as the prefix, but do not carry the associated meaning, e.g. *mis-times* (prefix) vs. *mistakes* (pseudo-prefix). Here, under certain linguistic frameworks, the morphological difference entails a different syllable structure: in the prefix, /s/ belongs to the coda of the first syllable and /t/ to the onset of the second syllable, while in the pseudo-prefix the /st/ cluster is ambisyllabic. Acoustic analysis revealed durational and spectral differences that are consistent with this analysis; for example, shorter /s/ in prefixes is indicative of its syllable-coda position, while more peripheral /ɪ/ in prefixes is likely to reflect the stronger rhythmic beat on the prefix due to the lack of a complex syllable coda. In this way, the phonetic exponents of syllable structure provide information about morphological structure.

Baker found that listeners are sensitive to these phonetic differences. In an intelligibility-in-noise task, listeners were more accurate in transcribing sentences with matching phonetic detail (created by splicing together two tokens of the same sentence) than those with inappropriate phonetic detail (created by splicing a prefix into a pseudo-prefix context or vice versa). Therefore, phonetic detail driven by differences in morphological structure has a role in perception.

This conclusion is further supported by Kemps et al. (2005a), who showed that Dutch listeners are sensitive to phonetic differences between inflected and uninflected forms. When inappropriate phonetic detail was introduced by cross-splicing a stem from a singular token (e.g., [buk] *boek* (*book*)) with the suffix from a plural token (e.g., [ən] from *boek-en* (*books*)), reaction times in a 'singular' vs. 'plural' decision task were slower than in a matched-splice control condition. In addition, using a lexical decision

task in both English and Dutch, Kemps et al. (2005b) found that listeners had longer reaction times when the monosyllabic stems of agent nouns and comparatives were spliced from their bisyllabic context (e.g., *work* from *worker*; *wet* from *wetter*) than when they were produced as bare stems. These studies indicate that listeners can use the phonetic detail that differs between bare stems and suffixed forms to inform metalinguistic judgements. Unlike Baker (2008), these studies did not control the number of syllables, and therefore listeners' judgements may reflect inferences about the number of syllables that are indirectly informative about the morphological structure.

In view of this evidence for the perceptual relevance of phonetic detail relating to morphological structure, Experiment 3 explored whether listeners can adapt to an atypical pronunciation that is systematically associated with an affix.

4.1.2 Methodological background

Experiment 3 investigates whether listeners can adapt to an atypical pronunciation of the English prefix *re-*. This prefix means 'again' (Chambers Dictionary, 1998); for example, *re-paint* means to paint something again. The citation-form pronunciation of *re-* is almost always [ri:] (*The English Pronouncing Dictionary*, 1997). Experiment 3 investigates whether listeners can adapt to a talker who pronounces /ri:/ as [rɪ] in prefixes, e.g., *re-think* is pronounced [rɪθɪŋk].

The [rɪ] pronunciation was chosen because the change from [ri:] to [rɪ] crosses a phoneme boundary and should therefore be extremely salient to listeners, but, at the same time, it remains a reasonably natural-sounding variant of /ri:/. For example, there are a few exceptional cases where both [ri:] and [rɪ] are acceptable realisations of the initial syllable /ri:/, in words such as *receptivity, recoup* and *repaid* (*The English Pronouncing Dictionary*, 1997). In addition, /i:/ may be produced less peripherally in rapid or casual speech.

The independent meaning conveyed by the *re-* prefix may make it less susceptible to reduction than the initial /ri:/ syllable in polysyllabic content words, in which the meaning is carried by several syllables. This would be consistent with the idea that words (or syllables) with a greater information load are less likely to be reduced. (This idea is more commonly expressed in terms of 'probability' (e.g., Jurafsky, 2001)

'redundancy' (Hunnicutt, 1985; Lieberman, 1963), frequency (e.g., Bybee, 1996) or novelty (Fowler & Housum , 1987)). Nevertheless, there are occasional instances of *re-* prefix reduction when the information load is particularly low. For example, a [rɪ] pronunciation seems common in cases where the morphological status of the 'prefix' is marginal due to historical vocabulary loss such that the bare stem is extremely infrequent or extinct (e.g., *recoup*, *regurgitate, resuscitate, rejuvenate*). In addition, informal personal observation suggests that a [rɪ] pronunciation of the prefix *re-* is occasionally produced in repeated words, i.e. where the information is 'given' rather than new. However, such examples are relatively rare, so there is no evidence to suggest that the [rɪ] pronunciation has a strong prior association specifically with the prefix *re-*.

Experiment 3 followed the same training-test paradigm as Experiments 1 and 2. In the training task, two groups of subjects were familiarised with the talker by listening to stories including either the atypical [rɪ] or the usual [ri:] pronunciation of the *re-* prefix (**Accent subjects** and **Control subjects** respectively).

An intelligibility-in-noise task was used to assess perceptual learning. A lexical decision task was not considered a useful assessment tool for the reasons discussed in Section 2.4.5 in relation to Experiment 1. Although a sentence completion task was successful in Experiment 2a, this task was not appropriate in Experiment 3 because of the nature of the stimuli. That is, assessing the perceptual learning effect in Experiment 3 requires detection of between-group differences in subjects' responses to the syllable /ri:/ pronounced as [rɪ]. Crucially, in Experiment 3, there are two valid interpretations of [rɪ] for Accent subjects: as an atypical pronunciation of /ri:/ in prefixes, or as a word beginning with the phoneme sequence /rɪ/ (e.g., *reply*, *receive*). If Accent subjects completed sentence fragments ending in word-initial [rɪ] on the latter basis, it would render the sentence completion task useless for assessing perceptual learning because their responses would be indistinguishable from those of Control subjects.

An intelligibility-in-noise task is an appropriate alternative for several reasons. Firstly, it is a relatively ecologically-valid task. It mimics an everyday situation – speech can

be masked by noise from crowded rooms or loud vehicles, etc. – and, as in ordinary conversation, the aim is to understand the meaning of the speaker.

Secondly, it enables the atypical pronunciation to be retained in its lexical and sentential context, which enhances naturalness and eliminates potential difficulties with truncation.

Thirdly, it is important to avoid the test task being too informative about the atypical accent, in order to minimise the adaptation of Control subjects and increase the chance of detecting between-group differences. In an intelligibility-in-noise task, if the signal-to-noise ratio (SNR) is low enough to make intelligibility difficult, exposure to /ri:/ realised as [rɪ] should not act as an effective learning experience for the Control group because the top-down lexical, semantic and syntactic information used to guide learning should be less reliable than that available in good listening conditions.

To enhance the naturalness of the task, multi-talker babble was used rather than white noise. Baker (2008) demonstrated listener sensitivity to phonetic detail in prefixes by using cafeteria noise to mask sentences. Multi-talker babble rather than cafeteria noise was chosen for this experiment because (when enough speakers are used) it is likely to be more spectrally-consistent than cafeteria noise. This should reduce the influence of windows of intelligibility or 'glimpses' (Cooke, 2003) on the intelligibility of particular words in the sentences, and so reduce the disproportionate influence of those words on subjects' interpretation of what they hear.

4.2 METHOD

4.2.1 Subjects

112 subjects (34 male, 78 female) participated in the experiment. Subjects were aged between 18 and 31, and were non-phonetically-trained native speakers of British English, monolingual (until at least age 11) with no known speech or hearing problems. They were nearly all students at the University of Cambridge. They were all unfamiliar with the voice of the talker in the experiment, and were paid £6 for their participation.

4.2.2 Materials

4.2.2.1 Training materials: Stories

The training task consisted of ten original short stories read by a male phonetician (aged 37) with an SSBE accent. Recording equipment was as described in Section 3.2.2.1.

Two versions of each story were recorded. In the Accent version, all instances of *re-* prefixes were realised as [rɪ]; and in the Control version, all instances of *re-* prefixes were realised as [ri:]. Accent and Control versions were spoken (as far as possible) at the same impressionistic rate and with the same intonation pattern. The recordings were checked auditorily by the author, and any divergent intonation patterns or realisations that were not sufficiently clear tokens of [rɪ] were re-recorded. The total duration of the stories was 18 m 59 s (Accent version) and 18 m 42 s (Control version). Appendix K contains the transcript.

The stories included 123 *re-* prefixes and 176 instances of /ri:/ in stems. Prefixes met the following criteria (after Wurm, 1997, p. 442): the meaning of the full form is semantically transparent, a free-standing word remains after the prefix is removed, and the meaning of the prefix is consistent with other instances of that prefix.

Steps were taken to optimise the chance of inducing morphologically-sensitive perceptual learning in Accent subjects. Firstly, any other affixes containing /ri:/ (e.g., *pre-*) were excluded. Secondly, *re-* prefixes followed by vowels were excluded because a palatal approximant is often inserted between the prefix and the stem (e.g., *re-animate, re-unite*), and this palatalisation diminishes the contrast between [ri:] and [rɪ]. Thirdly, particular effort was taken to include *re-* prefixes followed by unstressed syllables (e.g., *rediscover*) in the stories. This aimed to maximise the similarity of the prosodic context of *re-* prefixes between the training and test tasks. Word-initial /ri:/ in stems usually occurs before unstressed syllables (e.g., *reason*, *recent*), and so, due to the prosodic matching procedure described in Section 4.2.2.2, it is the context in which *re-* prefixes occurred most frequently in the intelligibility-in-noise task. *Re-* prefixes in the stories were followed by an unstressed syllable in 20% of tokens. Finally, /ri:/ in stems was included in various word positions, with and without

complex syllable onsets, in order to make clear that the atypical pronunciation was specific to prefixes.

To ensure subjects' attention, one comprehension question was included after each story (10 in total, listed in Appendix L).

4.2.2.2 Test materials: Intelligibility-in-noise sentences

Accent and Control subjects completed the same intelligibility-in-noise task, consisting of 18 experimental sentences and 60 filler sentences. These were recorded by the same talker and with the same recording equipment as the stories.

Each experimental sentence contained one instance of /ri:/ realised as [rɪ], either in a prefix (9 sentences) or in a stem (9 sentences). The words containing [rɪ] will be referred to as **keywords**. Keywords were novel (i.e. they had not occurred in the training phase). They were selected such that pronouncing /ri:/ as [rɪ] did not create a different word (e.g., *regal* would become *wriggle*). In addition, *re-* prefixes are, by definition, always in word-initial position, and therefore, in order to avoid a confound with syllable position, only stems with word-initial /ri:/ were used as keywords.

Prosodic differences between /ri:/ in stems and /ri:/ in prefixes were minimised by creating stem-prefix sentence pairs that shared many segmental properties and were matched in foot structure, nuclear stress position and F0 contour. Table 5 lists the 9 sentence pairs that were included in the experiment.

Filler sentences did not include the syllable /ri:/. 20 words (including 11 verbs) beginning with /r/ were included in the fillers in order to limit the usefulness of initial /r/ as a predictor of a *re-* prefix. Fillers are listed in Appendix M.

The babble used to mask the sentences was created from 12 British English speakers (4 male, 8 female) reading from a book or newspaper. It was recorded and mixed by Dr Antje Heinrich of Cambridge University Phonetics Laboratory. Different SNRs (signal-to-noise ratios) were created by holding the level of the noise (babble) constant, and adjusting the average amplitude of the sentences. To avoid an abrupt onset and offset, for each stimulus the babble gradually increased in amplitude for the first 200 ms of the sound file and decreased in amplitude for the final 200 ms. These 200 ms periods did not overlap with the target sentence.

Prefix sentences	**SNR (dB)**	**Keyword correct (%) from pilot data**	**Stem sentences**	**SNR (dB)**	**Keyword correct (%) from pilot data**
He aimed to **re-supply** the cocaine by Tuesday	2	20	They claimed the **recent** violent campaign was stupid	-5	33
The workers **re-secured** the entrance to the buildings	-3	63	These worthless **reasons** formed the basis of his dealings	1	30
We'll make his brother **re-consider** moving out of Woking	-3	55	They say this other **region's** an amazing place for walking	-9	57
He's been asked for a speedy **re-publication**	-6	58	She's been told it's a nasty **renal** infection	-1	30
They were happy to **re-connect** him to the digital archive	-5	53	They were better at **reaching** victims who were close to the stairwell	-3	53
He wants permission to **re-take** the photos	-8	51	He needs to visit the **retailer's** warehouse	-8	40
They were meaning to **re-display** them	-4	63	They were buying some **Reebok** trainers	-5	63
The marketing team had **re-defined** the brand	-6	55	The Corsican Pine in **Regent's** Park was brown	-7	60
Tom was certainly anxious to **re-submit** it	-3	55	John had told him to study the **rhesus** monkey	0	45

Table 5. Experimental sentences (Prefix and Stem) for the intelligibility-in-noise task, alongside the SNR at which each sentence was presented and the percentage of keywords correct in the pilot experiments at that SNR. Keywords are in bold.

To ensure keyword (e.g., *re-supply*, *recent*) intelligibility between 20% and 70%, the SNR was determined separately for each experimental sentence through piloting. This intelligibility range aimed to make sure that the sentences were possible to understand (avoiding a 'floor effect') while leaving 'room for improvement' due to training (avoiding a 'ceiling effect').

The most appropriate SNR for each sentence was established iteratively through a series of six pilot experiments, involving 58 subjects in total. To ensure talker-familiarity, pilot subjects also completed the training phase (Control version) of the experiment and the intelligibility-in-noise fillers. As these scores were from subjects who heard the Control version of the stories, they should represent the lower bound of intelligibility scores. Nine pairs of sentences were identified for use in the main experiment. The SNR that produced an appropriate intelligibility level for each sentence is listed in Table 5, alongside the keyword intelligibility for the sentence at that SNR.

The SNR for fillers ranged from -10 dB to -1 dB. This is lower than the SNR range for the experimental sentences (-9 dB to 2 dB) because pilot experiments using the same range as experimental sentences gave a higher overall intelligibility for fillers, presumably because experimental sentences contain atypical pronunciations, while filler sentences do not.

4.2.3 Procedure

Testing conditions, equipment and presentation software were identical to Experiments 1 and 2. Subjects were presented with written instructions which were clarified by the experimenter[6] as necessary.

Subjects were assigned randomly to one of two Training groups. 56 subjects heard only the Accent version of the story and 56 subjects heard only the Control version of the story.

The experiment began with a short practice session for the intelligibility-in-noise task (5 sentences) which included a range of SNRs (-2 to -6 dB). The sentences were

[6] For 13 subjects, the experimenter was an undergraduate Linguistics student, rather than the author.

presented in order of decreasing SNR. The practice sentences did not include the syllable /ri:/. A list is in Appendix N.

To monitor learning, the experiment was divided into two halves. Each half consisted of training (5 stories, each followed by one comprehension question), followed by testing (38 or 40 intelligibility-in-noise sentences).

The instructions informed subjects that they should listen carefully to the 'stories', bearing in mind that they would be asked about them later. All subjects heard the stories in the same order.

For the intelligibility-in-noise task, subjects were instructed to type what they thought they heard. If they were not sure, they should type their 'best guess'. 4 pairs of experimental sentences were presented within the first intelligibility-in-noise task, and 5 pairs within the second intelligibility-in-noise task (or vice versa). Experimental sentences were pseudo-randomised such that they were separated by at least one filler sentence. The order of items in the intelligibility-in-noise task was counterbalanced within Accent and Control Training groups. The experiment took approximately 50 minutes to complete.

4.3 HYPOTHESIS

Based on evidence from production and perception studies supporting an association between phonetic detail and morphological structure, it was predicted that listeners should adapt to a novel morphologically-dependent phonetic characteristic. Accent subjects were predicted to perform better than Control subjects for Prefix sentences in the intelligibility-in-noise task, and the same as Control subjects for Stem sentences.

The expected rate of learning was uncertain. If learning occurred between the First and Second halves of the experiment, responses were predicted to become more consistent with the training accent in the Second half.

4.4 RESULTS

4.4.1 Data pre-processing

Ambiguous responses (0.6% of the data) were excluded from the analysis. For example, it was unclear whether *risus* was an incorrect spelling of *rhesus* or an attempt to phonetically spell out the non-word [rɪsəs].

Subjects scored an average of 82% on the comprehension questions, ranging from 30%–100% correct. Most of the incorrect responses still made reference to aspects of the stories and therefore also indicate that subjects were attending to the content of the stories.

4.4.2 Analysis of intelligibility-in-noise task

Three different scoring methods were used to measure intelligibility:

Keyword correct: responses coded as correct only if the keyword was absolutely correct (or obviously correct, but spelled wrongly).

Keyword or suffixed variant correct: responses coded as correct if the keyword was absolutely correct (or misspelled) or a suffixed variant of the keyword, e.g. *re-defining* for *re-defined*.

Morphological type correct: responses coded as correct if [rɪ] was interpreted as a /ri:/ syllable of the correct morphological type (i.e. as a prefix or as part of stem), e.g. *re-commit* for *re-connect* or *reason* for *region*.

A mixed-effects model was fitted separately to the data from each scoring method. Each mixed model had fixed factors of Training group (Accent, Control), Morphological type (Prefix, Stem), Half (First, Second) and all interactions, and random factors of Subject and Item. The analyses were conducted using R version 2.6.0 (R Development Core Team, 2007).

4.4.2.1 Keyword correct

The most parsimonious model using *keyword correct* data included a significant interaction between Training group and Morphological type ($\chi^2 = 4.7$, $df = 1$, $p < 0.04$) and between Morphological type and Half ($\chi^2 = 4.1$, $df = 1$, $p < 0.04$).

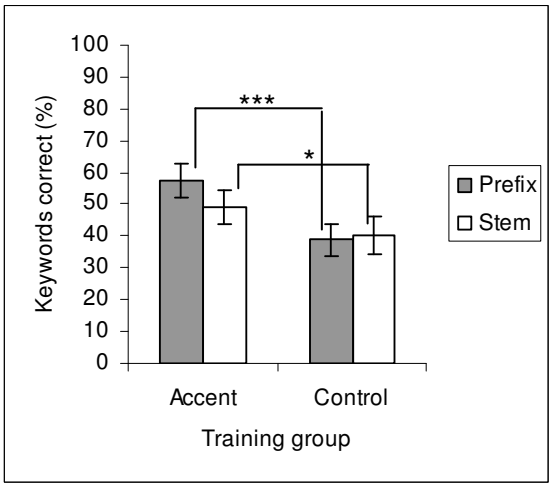


Figure 9. Percentage of keywords correct in the intelligibility-in-noise task using the *keyword correct* criterion, as a function of Training group and Morphological type. Error bars show the 95% confidence interval of the mean. Significant pairwise comparisons are indicated as follows: * $p < 0.05$, *** $p < 0.001$.

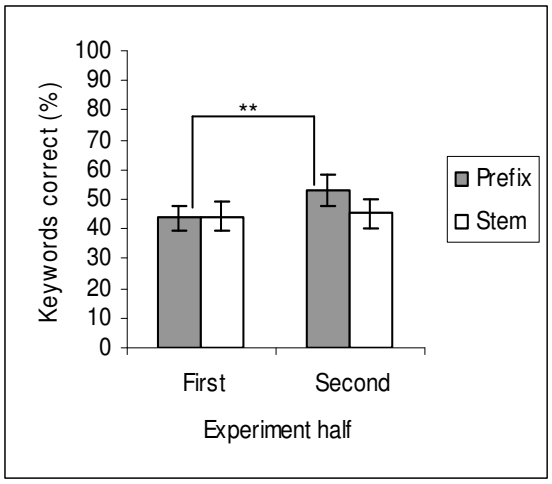


Figure 10. Percentage of keywords correct in the intelligibility-in-noise task using the *keyword correct* criterion, as a function of Half and Morphological type. Error bars show the 95% confidence interval of the mean. Significant pairwise comparisons are indicated as follows: ** $p < 0.01$.

Figures 9 and 10 show the results of pairwise comparisons exploring these interactions. Accent subjects had significantly higher keyword intelligibility scores than Control subjects for both prefixed words (57% vs. 39%, $z = 4.9$, $p < 0.001$) and stems (49% vs. 40%, $z = 2.4$, $p < 0.02$), but this difference was significantly greater for prefixed words. The better performance of the Accent subjects across both prefixed words and stems indicates that perceptual learning generalised across Morphological Type, while the advantage for prefixed words demonstrates some specificity of perceptual learning in relation to morphological structure.

Further evidence for the specificity of perceptual learning was found in the second half of the experiment. Pairwise comparisons showed that Accent subjects had significantly higher keyword intelligibility scores for prefixed words than stems (62% vs. 51%, $z = 2.3$, $p < 0.03$) in the second half, while the difference for Control subjects

was not significant (44% vs. 39%, $z = 0.6$, $p = 0.6$). However, this between-group difference was not sufficient to produce a significant three-way interaction between Training group, Morphological type and Half.

The change in keyword intelligibility between the first and second halves of the experiment was dependent on Morphological type. Both Accent subjects and Control subjects found prefixed words to be significantly more intelligible in the second half of the experiment than the first half (53% in the first half vs. 62% in the second half for Accent subjects, 34% vs. 44% for Controls; $z = 3.1$, $p < 0.003$), but there was no significant difference in intelligibility for stems ($z = 0.2$, $p < 0.9$). This improvement in intelligibility for prefixed words over the course of the experiment was expected for Accent subjects, but not for Control subjects.

4.4.2.2 Keyword or suffixed variant correct

Using the *keyword or suffixed variant correct* criterion, evidence for the generalisation of perceptual learning remained robust; there was a significant main effect of Training group, with Accent subjects scoring significantly better than Control subjects (55% vs. 43%; $\chi^2 = 14.9$, $df = 1$, $p < 0.001$). However, contrary to the prediction of perceptual learning associated with prefixes for Accent subjects, the interaction between Training group and Morphological type was no longer significant, although the trend was in the predicted direction (Accent subjects: 61% for prefixed words and 50% for stems; Control subjects: 45% for prefixed words and 40% for stems). The interaction between Morphological type and Half was only marginally significant ($\chi^2 = 3.3$, $df = 1$, $p < 0.07$).

4.4.2.3 Morphological type correct

Using the *morphological type correct* criterion, the pattern of results was similar to the *keyword correct* criterion, with the exception that the interaction between Training group and Morphological type was only marginally significant ($\chi^2 = 3.1$, $df = 1$, $p < 0.08$).

4.4.3 Filler sentences in intelligibility-in-noise task

To assess whether there were any differences in task performance between Training groups, an analysis of subjects' responses to the fillers in the intelligibility-in-noise task was also conducted.

The intelligibility of filler sentences was assessed by scoring a **target word** within each sentence as correct or incorrect. Target words were content words, usually polysyllabic, chosen from positions towards the middle and end of sentences, reflecting the position of keywords within the experimental sentences (see Appendix M). The accuracy of target words was scored using two methods: the *target word correct* criterion, in which the word had to be absolutely correct (or obviously misspelled); or the *target word or affixed variant correct* criterion. These gave the same basic pattern of results, but with intelligibility increased by approximately 2% for the *target word or affixed variant correct* criterion, so only the results from the *target word correct* analysis are reported.

4.4.3.1 Target word correct

A mixed model analysis was carried out using *target word correct* data with fixed factors of Training group, Half and Training group × Half and random factors of Subject and Item. There was a significant interaction between Training group and Half ($\chi^2 = 5.1$, $df = 1$, $p < 0.03$). Pairwise comparisons (see Figure 11) show that Control subjects scored more highly in the second half of the experiment than in the first half (67% in the first half vs. 73% in the second half; $z = 4.7$, $p < 0.001$), but the change was not significant for Accent subjects (70% vs. 71%; $z = 1.5$, $p = 0.1$). There was no significant effect of Training group in either Half.

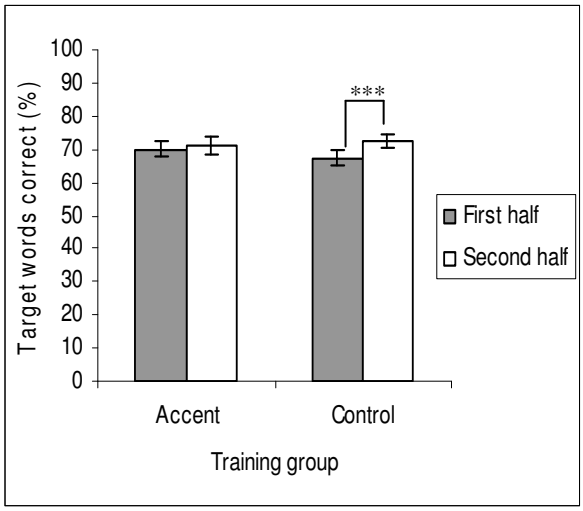


Figure 11. Percentage of target words correct in the intelligibility-in-noise task using the *target word correct* scoring system, as a function of Training group and Half. Error bars show the 95% confidence interval of the mean. Significant pairwise comparisons are indicated as follows: *** $p < 0.001$.

4.5 DISCUSSION

The results of Experiment 3 are complex, showing that perceptual learning is sensitive to morphological context yet also generalises from prefixes to stems. As discussed below (Sections 4.5.1 and 4.5.2), the co-existence of generalisation and specificity may indicate that the atypical realisation became associated with multiple levels of representation, or that a strategic response accompanied low-level category adaptation. The intelligibility gain for prefixed words for Control subjects between the first and second half of the experiment suggests there may be differences in the processing of atypical pronunciations depending on their morphological context (see Section 4.5.3).

4.5.1 Perceptual learning by Accent subjects is sensitive to morphological context

Accent subjects heard /ri:/ pronounced as [rɪ] in prefixes during the training phase, but had higher intelligibility scores than Control subjects for both prefixed words *and* stems, indicating that learning had generalised beyond prefixes. However, crucially, intelligibility advantage for Accent subjects was significantly greater for prefixes than stems (at least for the *keyword correct* scoring system), indicating that perceptual learning was to some extent specific to prefixes.

It is unlikely that the prefix-specific learning effect can be explained in terms of (non-phonetic) expectations raised by the frequent occurrence of atypically-pronounced *re-* prefixes during the training phase. The frequency of *re-* prefixes was the same for Accent and Control subjects, but the atypical pronunciation for Accent subjects may have drawn their attention to the *re-* prefixes. This could potentially have led Accent subjects to include more *re-* words in their responses to the intelligibility-in-noise task, leading to more correct responses for the prefixed keywords. However, an examination of filler responses suggests this did not occur. A relatively large proportion of fillers were used in the assessment task in order to minimise the possibility of using expectations of a *re-* prefix as a guide. Almost a third of fillers included word-initial /rV/, but only four filler responses (0.0006%) included a *re-* prefix; two of these were from Accent subjects and two from Control subjects. This shows that Accent subjects' intelligibility advantage for prefixed words does not result from a tendency to guess at *re-* prefix words without good auditory evidence. It

suggests that their performance reflects probabilistic interpretation on the basis of the auditory signal and their knowledge of the talker's atypical 'accent', rather than more random guesswork based on a notion that the experiment is investigating *re-* prefixes.

Evidence for morphologically-sensitive perceptual learning supports a model of speech perception in which the speech signal is interpreted in the context of morphological structure. In this way, new associations between phonetic characteristics and morphological structure can be learned, and can then influence the perception of morphological structure. Perceptual learning solely in terms of phonemes or features cannot account for morphologically-sensitive learning.

4.5.1.1 Differences in scoring methods

The interaction between Training group and Morphological type differed in significance across the different scoring methods. This suggests that the effect of (un)familiarity with the atypical pronunciation was quite subtle.

The *keyword correct* criterion showed a significant interaction between Training group and Morphological type, while the *Keyword or suffixed variant correct* criterion did not. This implies that an important contribution to the interaction was made by the keywords for which Control subjects were able to identify the *re-* prefix but made an error in identifying the suffix. It is likely that, because the atypical pronunciation was consistent with Accent subjects' experience during the training phase, they were able to process the prefix more efficiently than Control subjects, and so were able to give more attention to the end of the keyword. So, pronouncing /ri:/ as [rɪ] in prefixes in the intelligibility-in-noise task did not necessarily *prevent* identification of the *re-* prefix for Control subjects, but perception of the *re-* prefix was *disrupted* and this had a detrimental effect on perception of the suffix.

The *Morphological type correct* criterion showed a marginally significant interaction between Training group and Morphological type. This scoring method entailed scoring all suffixed variants as correct. Therefore, the critical items which distinguish this scoring system from the *suffixed variant* scoring system are those in which [rɪ] was interpreted as /ri:/ but the rest of the word was wrong (e.g., *region* for *reason*, or *re-commit* for *re-connect*). Consequently, the (marginal) re-emergence of the greater Prefix advantage for Accent subjects shows that, when most of the keyword was

incorrectly perceived, [rɪ] was interpreted as /ri:/ more often in the context of prefixed keywords than stem keywords for Accent subjects. This suggests that when the identification of keywords was very difficult, Accent subjects' familiarity with the atypical pronunciation in prefixes increased the likelihood that they would interpret the initial syllable as a *re-* prefix. Therefore, familiarity with the way a talker produces a prefix can influence perception in difficult listening conditions.

Despite recent research on perceptual learning, little is known about the role that talker-familiarity plays in speech perception – under what conditions is it useful, and what effects does it have in natural listening situations? The different results from the scoring methods in Experiment 3 suggest that familiarity with the way a talker makes morphological distinctions can have different effects depending on the precise listening situation; where perception is very difficult it can influence the assignment of morphological structure, while in slightly less difficult circumstances it may affect the speed or efficiency of processing. These conclusions can only be tentative on the basis of the present results, but nevertheless they suggest that there is considerable scope for research into the extent to which listeners use newly-acquired knowledge of a talker in various listening contexts.

As well as being of theoretical interest, a better understanding of when and how talker-specific knowledge can influence perception has implications for methodology. For example, in this experiment, familiarity with the way a talker produces prefixes seemed to have a positive effect on the perception of suffixes in the same word. This suggests that perceptual learning studies should focus not only on the interpretation of particular phonetic cues, but on possible effects on attention that may affect processing later in an utterance.

4.5.2 Co-existence of context-sensitivity and generalisation

Section 4.5.1 focussed on the interaction between Training group and Morphological type, which supported the hypothesis that perceptual learning is sensitive to morphological structure. However, contrary to the hypothesis, there was also very robust evidence of generalisation. Accent subjects scored significantly better than Control subjects for both prefixed words and stems in the intelligibility-in-noise task, under all scoring methods.

It is unlikely that the existence of both generalisation and context-sensitive perceptual learning represents two groups of Accent subjects learning different things. There were relatively few responses per subject, so it is difficult to draw firm conclusions about each individual subject's behaviour. Nevertheless, for each subject, a higher number of correct Prefix responses than correct Stem responses indicates a higher probability that perceptual learning was specific to prefixes. Figure 12 shows the distribution of these scores for Accent subjects. It shows a clear unimodal distribution; there is no obvious clustering of subjects into a group which find prefixes more intelligible, and a group which finds prefixes and stems equally intelligible. Therefore, it is likely that generalisation *and* specificity are present within individual subjects.

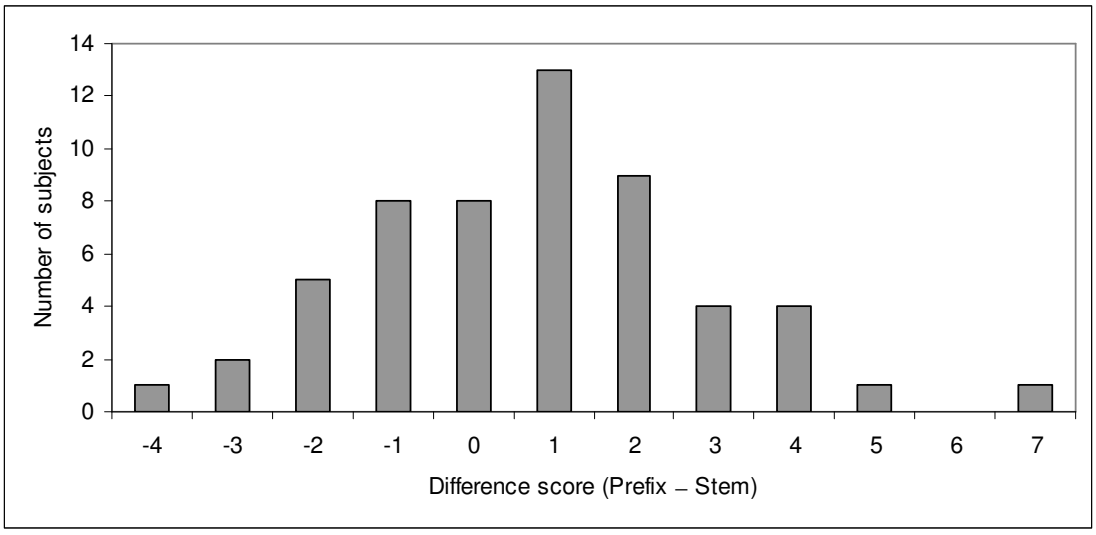


Figure 12. The number of Accent subjects achieving each difference score is shown. The difference score is the number of correct Prefix responses minus the number of correct Stem responses in the intelligibility-in-noise task, using the *keyword correct* criterion. A positive score indicates that subjects found prefixed words more intelligible, i.e. it indicates a higher probability of prefix-specific perceptual learning.

There are several ways in which both prefix-specific learning and generalisation to stems could occur at the same time, depending principally on the mechanisms underlying generalisation. Generalisation may be due to concurrent adaptation of categories at different levels of structure, or it may be the result of a task-related strategy such as a more tolerant approach to atypical pronunciations. These possibilities are discussed below.

4.5.2.1 Generalisation due to adaptation of category representations at multiple levels of structure

Concurrent adaptation of categories at many different levels of structure is plausible within a polysystemic approach. If a phonetic characteristic is processed within the context of all kinds of structures and systems, it seems likely that associations of varying strengths could be learned at several levels of structure at the same time.

Accent subjects learned to associate the [ɪ] realisation with the *re-* prefix, but they may also have learned a weaker association with a broader category or categories such as the syllable /ri:/, and/or the vowel /i:/. Listeners could also have learned about the talker's rhythm (e.g., that the talker may shorten certain segments). In addition, learning may or may not have been related to word position (all the atypical pronunciations occurred in word-initial syllables) and may or not apply to syllables with complex syllable onsets (e.g., *street*).

These possibilities could be investigated using a series of experiments that adapt the intelligibility-in-noise task to include atypical realisations in a range of contexts. For example, /ri:/ could be realised as [rɪ] in complex-onset syllables or non-word-initial positions; the vowel /i:/ could be realised as [ɪ] in various phonetic contexts; or a durational manipulation that reflects the [i:]–[ɪ] shift could be incorporated. In this way, the locus (or loci) of generalised perceptual learning should become clearer.

Concurrent adaptation at multiple levels of structure is discussed further in Chapter 5.

4.5.2.2 Generalisation due to task-level responses

Alternatively, higher intelligibility scores for Accent subjects could reflect a different approach to the intelligibility-in-noise task, as compared to Control subjects. For example, the atypical pronunciation during the story could have caused Accent subjects to focus on the phonetic detail of the speech signal, which may be a more effective strategy for the intelligibility-in-noise task. The analysis of filler responses (see Section 4.4.3) suggests this is unlikely. Higher scores for fillers for Accent subjects would have indicated that the 'generalisation' was a consequence of a task-related strategy that improved overall performance. However, although there was an interaction between Training group and Half, there was no difference in intelligibility scores for fillers between Accent and Control subjects in either Half. This suggests that higher intelligibility scores for Accent subjects for prefixed words and stems are

not due to a strategy which improves performance on the intelligibility-in-noise task as a whole. Instead, the intelligibility advantage for Accent subjects is restricted to the atypically-pronounced keywords.

One strategy that may improve performance only for atypically-pronounced keywords is to adopt a 'tolerant' approach to atypical pronunciations. That is, Accent subjects may be more flexible in their interpretation of *any* atypical pronunciation. In other words, they are more likely to type a response even when they are aware that what they have written does not fit the typical realisation of that word. Because there *are* atypical pronunciations within the intelligibility-in-noise task, this strategy could be effective. On the other hand, it is arguable that a strategy which essentially amounts a greater propensity to guess would not necessarily improve subjects' performance; in order to be effective, it is also necessary that subjects' guesses should be correct.

Experiment 3 cannot resolve whether generalisation was due subjects adopting a 'tolerant' approach, or to subjects learning a specific association between the [ɪ] and several categories or levels of structure. However, the issue could be addressed by including different atypical realisations in the assessment task from those used in the training task (cf. Maye et al., 2008). For example, the intelligibility-in-noise sentences could include /ri:/ pronounced as [rɛ] or [rʌ] as well as [rɪ]. A 'tolerant' strategy predicts that Accent listeners should find all atypical pronunciations more intelligible than Control listeners. Learning of a specific association between [ɪ] and /ri:/ or /i:/ predicts that only /ri:/ pronounced as [rɪ] should be more intelligible.

4.5.3 The influence of Morphological type on perceptual learning by Control subjects

The interaction between Morphological type and Half shows that improvement in intelligibility scores between the first and second halves of the experiment is dependent on Morphological Type. Both Accent subjects and Control subjects found significantly more prefixed words intelligible in the second half of the experiment, but there was no significant change for stems. This prefix-specific learning effect was expected for Accent subjects, who heard the atypical pronunciation in prefixes during the training task and were therefore likely to adapt over the two halves of the experiment. However, it was not predicted for Control subjects.

Control subjects heard /ri:/ pronounced as [rɪ] in both prefixes and stems during the intelligibility-in-noise task, but only the intelligibility of prefixed words improved between the first and second half (an increase of 10% for prefixed words, and a decrease of 2% for stems, using the *keyword correct* criterion). The analysis of fillers indicates that increased intelligibility is not exclusive to prefixed words; Control subjects also showed significantly higher target word intelligibility for fillers in the second half of the experiment (an increase of 5% using the *target word correct* criterion). Therefore, the intelligibility gain may be partially due to a task-level process.

For example, increased attention to phonetic detail may be responsible for Control subjects' higher scores in the second half of the experiment. The intelligibility-in-noise task was used in the hope that it would provide Control subjects with relatively little useful top-down information concerning the atypical pronunciations, and therefore perceptual learning during assessment would be unlikely (see Section 4.1.2). Nevertheless, those experimental sentences with a particularly high SNR contained relatively clear atypical pronunciations which could have acted as training. If Control subjects became aware during the first half of the assessment task that some words sounded atypical, this would have been likely to focus their attention on the phonetic detail of the signal, leading to an improvement in performance for fillers. In contrast, Accent subjects are likely to have paid close attention to the phonetic detail from the beginning of the assessment task, having noted the atypical realisations in the stories, and so do not show the equivalent improvement between the first and second half for fillers.

However, Control subjects' performance cannot be explained solely by increased attention to the phonetic detail of the speech signal because the improvement in intelligibility was greater for prefixed words than for fillers, and performance decreased for stems. Furthermore, although greater attention to phonetic characteristics may account for the increase in filler intelligibility, it will not necessarily facilitate the perception of atypical realisations unless it is accompanied by appropriate knowledge. This suggests that Control listeners learned the atypical pronunciation in prefixes, but not in stems, despite being exposed to both of them in

the assessment task. Therefore, the atypical pronunciation seems to be more learnable when it occurs in a *re-* prefix than a stem.

The possibility that the intelligibility-in-noise task provided more top-down guidance for prefixed words is unlikely because there was no significant difference in the overall intelligibility of prefixed words or stems in the first half of the experiment (for Control subjects, prefixes were actually 7% less intelligible using the *keyword correct* criterion).

It is likely that (other things being equal) phonetic characteristics that relate to a particular semantic distinction (such as the 'again' meaning of the *re-* prefix) are easier to learn than when same phonetic characteristic is not associated with a such a clear-cut semantic distinction. This is based on the assumption that the systematic co-occurrence of a salient phonetic characteristic and any pre-existing category will result in learning an association between the two. The more attributes that consistently 'define' the pre-existing category, the more 'robust' the category, and the easier the learning task. Thus, the atypical pronunciation of /ri:/ as [rɪ] will be more easily related to the *re-* prefix than /ri:/ in stems because it is a consistent semantic entity, as well as a phonetic/phonological one.

4.6 CONCLUSION

Accent subjects in Experiment 3 demonstrated perceptual learning that was conditioned by morphological context, providing further evidence that listeners can adapt linguistic categories that relate to higher levels of linguistic structure. The co-existence of generalisation and specificity may provide evidence that not only is perceptual learning context-sensitive, but it is sensitive to many different contexts at the same time.

Control subjects showed morphologically-sensitive perceptual learning that indicated that the atypical pronunciation was easier to learn in prefixes than stems. Like Experiment 2, this provides evidence that some associations are easier to learn than others. However, in this case, the explanation may relate to the robustness of the prefix category, rather than the existence of a prior association between the atypical variant and the prefix category.

CHAPTER 5

GENERAL DISCUSSION

5.1 INTRODUCTION

As reviewed in Chapter 1, there has been little previous research investigating the plasticity of phonetic categories with reference to categories or contexts above the level of features or phonemes. This thesis investigated perceptual learning within a framework of speech perception that acknowledges the importance of phonetic detail at multiple levels of representation. The results highlight the importance of the linguistic-structural context in which an atypical variant occurs, and demonstrate that listeners can adapt to phonetic detail that is systematically associated with prosodic and morphological structure. The experiments also begin to shed light on the constraints that govern perceptual learning by showing that some associations are more learnable than others. This helps to elucidate the balance between the stability and plasticity of the perceptual system as it adapts to novel information.

The theoretical implications of the present research are discussed in more detail in Sections 5.2–5.4 in the context of proposed properties of a model of perceptual learning, while the methodological implications are discussed in Section 5.5.

5.2 PROPOSED PROPERTIES OF A MODEL OF PERCEPTUAL LEARNING

This section outlines the proposed properties of a model of perceptual learning. It is based on a polysystemic approach to speech perception, as described in Polysp (Hawkins & Smith, 2001), in which properties of the acoustic signal are assumed to be relevant to multiple categories (linguistic and non-linguistic), and the speech signal is processed and understood in the context of multiple levels of representation (see Chapter 1 for evidence supporting this approach).

In Polysp, listeners' adaptation follows as a consequence of self-organising categories that emerge as a result of the distributional regularities of the input, modulated by attention and task (Hawkins & Smith, 2001, p. 138). However, Polysp lacks a detailed

and explicit account of the influences on perceptual learning. The approach taken in this chapter is perhaps more strongly exemplar-based than the Polysp approach, which argues that 'both exemplar representation and abstraction are necessary' (Hawkins, 2010, p. 484). In the present account, representations of categories and structures are taken to emerge from generalisations across previously-learned exemplars.

Many exemplar accounts are based on evidence for the retention of phonetic detail that relates to syntactic information (Hay & Bresnan, 2006) or to social information such as gender, social class, age and regional background (Hay et al., 2006a; Hay et al., 2006b; Johnson et al., 1999), and therefore they try to explain how listeners have acquired particular context-sensitive phonetic categories. Others are motivated by talker-familiarity effects (Goldinger, 1996, 1998; Lachs et al., 2003; Nygaard & Pisoni, 1998), and so discuss evidence that the perceptual system adapts to novel information, but they are not specific about the phonetic characteristics that are learned. The present account brings together the exemplar approach and evidence about the phonetic characteristics that listeners can learn most easily from exposure to a talker, and places these in the context of a polysystemic approach to speech perception in which phonetic properties are assumed to be relevant at multiple levels of representation.

A detailed implementation of a model of perceptual learning at the neural level is not attempted here, and no computational implementation is offered. Rather, this section suggests general properties that such a model should include, based on the present experiments and the literature on perceptual learning. Further research is required to confirm and elucidate many details, and interdisciplinary expertise is required to produce a more computationally or neurally explicit model.

Before outlining the suggested properties of a model of perceptual learning, the terms 'phonetic property', 'category' and 'association strength' are defined.

'Phonetic properties' refer to speech-related input to the perceptual system (primarily auditory, but not excluding visual). In connectionist terms, the activation of phonetic properties might be seen as analogous to the values of feature dimensions that are used as input in a model such as TRACE (McClelland & Elman, 1986) – although, unlike in TRACE, this input can be affected by top-down influences (see Property 5,

below), resulting in bidirectional information flow throughout the perceptual system, rather than just in the upper layers as in TRACE. In neural terms, 'phonetic properties' refer to neurons (or, more likely, groups of neurons) that are activated at a very early stage in perceptual processing.

A 'category' refers to a previously-learned grouping of exemplars based on their mutual similarity. Categories can be linguistic (representing allophones, phonemes, words, lexical stress, prosodic position etc.) or non-linguistic (e.g., representing a particular talker, social group, or emotion). The activation of a category might correspond to the value of a unit at the feature-, phoneme- or word-level in TRACE. In neural terms, categories may be 'cell assemblies' (Hebb, 1949) or 'functional webs' (Pulvermüller, 2002), i.e. strongly-connected sets of neurons capable of working together.

The 'strength' of an association is defined as the extent to which a phonetic property is able to activate a category, and vice versa. 'Association strength' is analogous to 'connection weights' in TRACE.

The proposed properties are as follows:

1. Exemplars should be conceptualised as 'processing memories' rather than unanalysed episodes, i.e. each stored exemplar reflects how it was processed, and so it is affected by attention and task demands (cf. Pierrehumbert, 2006, p. 525). Exemplars, in a polysystemic account, are experiences that are processed in a context-sensitive way with reference to many linguistic and non-linguistic categories at the same time. Each stored exemplar is therefore a combination of perceptual input and prior knowledge (cf. Goldinger, 2007, p. 50).

2. An input (or exemplar) activates previous exemplars to the extent that they are similar to the input, and connections are strengthened between entities that are activated concurrently (as discussed in Polysp, this might be implemented neurally partly through Hebbian learning and cell assemblies/functional webs (Hebb, 1949; Pulvermüller, 2002)).

3. Listeners can learn associations of varying strengths between phonetic properties and many categories at multiple levels of representation.

4. The strength of a novel association is dependent on the degree of activation of the phonetic property and the degree of activation of the to-be-associated category (together with the strength of any pre-existing associations between the two). Higher short-term activation leads to stronger long-term associations (though repetition of experience is important for long-term adaptation – see Property 7).

5. The degree of activation of a phonetic property is determined by its perceptual salience in the current task. Perceptual salience is affected by the inherent auditory (or visual) salience of aspects of the signal, and by attention, informational value (determined by prior experience) and the degree to which the signal conforms to expectations.

6. The degree of activation of a category is affected by its robustness (i.e. the number and perhaps variety of attributes that co-occur to form this category, and the consistency with which they do so), by attention, and by the degree of activation triggered by bottom-up sensory information through prior learned associations.

7. The strength of an association between a phonetic property and a category is also dependent on the frequency and consistency with which the phonetic property is activated concurrently with the relevant category.

The following sections discuss these properties in relation to the present research.

5.3 ADAPTATION TO PHONETIC ASSOCIATIONS AT MULTIPLE LEVELS OF REPRESENTATION

A polysystemic exemplar-based approach (as summarised in Properties 1 and 2) predicts that listeners should be able to learn novel phonetic associations with multiple levels of linguistic structure (including above the level of the phoneme or the syllable it is in) or with social, pragmatic or other aspects of a communicative situation. The primary aim of the present research was to test the hypothesis that listeners can adapt to phonetic detail associated with higher levels of linguistic structure. As is described in this section, the results of all three experiments emphasised the importance of linguistic-structural context in perceptual learning, and Experiments 2 and 3 demonstrated perceptual learning of phonetic detail that is conditioned by linguistic-structural context. The results of Experiment 3 are also

discussed in relation to learning at multiple levels of structure concurrently, with reference to Properties 3 and 7.

Experiment 1 showed that the stress context of an atypical variant can influence perceptual learning. Global subjects, who heard /ɪ/ pronounced as [ə] in both stressed and unstressed syllables during training, were more likely to accept Stressed-vowel stimuli (e.g., *disco* pronounced as [ˈdəskəʊ]) as real words in a lexical decision task than Stress-conditioned subjects, who heard the same number of atypical variants in training, but only in unstressed syllables. This strongly suggests that the same phonemic change can have different effects depending on its stress context, which argues against modelling perceptual learning solely in phonemic or featural terms (cf. Kraljic & Samuel, 2006) because, by definition, phonemes exclude stress. Rather, this result supports an account in which an atypical variant is processed with reference to its prosodic context.

Although Stress-conditioned subjects did not seem to learn an association between centralisation of /ɪ/ and unstressed syllables, in that the predicted increase lexical acceptance rate for Unstressed-vowel stimuli did not occur, there was nevertheless tentative evidence to suggest that centralisation of /ɪ/ in unstressed syllables did influence Stress-conditioned subjects' responses. The analysis of data from Stressed-vowel stimuli in the lexical decision task showed that Stress-conditioned subjects did not adapt to the centralisation as Task number increased, whereas Control subjects showed a significant learning effect. It seems probable that Stress-conditioned subjects responded to the Stressed-vowel stimuli on the basis of their experience of the talker during training, while Control subjects, for whom the accents in training and at test were 'more different', dissociated training and test tasks (see Section 2.4.3.1). This suggests that an atypical segmental realisation in an unstressed context may produce subtle perceptual learning effects, perhaps influencing a general percept of talker-identity or -similarity instead of modifying lexical decisions to stimuli that include the unstressed variant. Again, this supports an approach which goes beyond a phoneme-, feature- or syllable-based model of perceptual learning, and takes prosodic context into account.

Experiments 2 and 3 show that not only is the degree of perceptual learning affected by the linguistic-structural context in which an atypical variant occurs (as shown in Experiment 1), but listeners can also learn to associate phonetic detail with higher levels of linguistic structure.

There is considerable evidence that prosodic cues, including duration, have a strong influence on speech perception (see Section 1.3.3), including at the level of phonological phrase boundaries (see Section 3.5). Experiment 2 demonstrates that listeners are able to adapt to duration differences that reflect prosodic structure, and can use their knowledge of a talker's accent to influence their interpretation of syntactic structure.

Experiment 2a was originally designed to induce perceptual learning of a durational cue associated differentially with nouns and verbs. However, because of the natural confound in English between nouns and syntactic phrase-final position, the duration manipulation was also correlated with prosodic phrase position. As a consequence, the perceptual learning effect present within the sentence completion task using phrases of the type *the Indian braves…* could have reflected learning of a direct association between the durational cue and grammatical category, or an association with phonological phrase position. Experiment 2b de-confounded the assessment task by using sentence beginnings of the type *she made [Joanna/Joanne a] dress so that…*, in which both noun and verb interpretations place the keyword (e.g., *dress*) in phrase-final position. The lack of a perceptual learning effect indicated that the difference between nouns and verbs observed in Experiment 2a was due to subjects learning an association between duration and phonological phrase position.

The duration manipulations in Experiment 2 did not cross a phoneme (or feature) boundary, and so the effect of duration on the interpretation of phonological phrase boundaries in Experiment 2a provides further evidence that listeners need not abstract away from the acoustic signal to features or phonemes at an early stage in perceptual processing. Furthermore, listeners' ability to learn associations with phonological phrase boundaries indicates that a dynamic model of speech perception must also include levels of structure above the word. Thus, an exemplar model based purely on an episodic lexicon (e.g., Goldinger, 1996, 1998) is unable to account for this result.

Instead, Experiment 2a shows that subtle duration differences can become associated with higher levels of linguistic structure, as is predicted by a polysystemic exemplar-based account. The reason that listeners learned an association with prosodic phrase boundaries rather than syntactic class may relate to Property 6, discussed in Section 5.4.2.2.

Experiment 3 provides further support for a polysystemic account of perceptual learning by showing that listeners are also able to adapt to phonetic detail associated with morphological structure. 'Accent' listeners, who were exposed to /ri:/ pronounced as [rɪ] in *re-* prefixes in the training story, demonstrated perceptual learning that was (at least partially) specific to *re-* prefixes in the intelligibility-in-noise task. Again, this demonstrates that a model of speech perception based on early 'matching' of particular realisations to phonemes or features without regard to linguistic-structural context is inadequate. In the right circumstances, listeners are able to use what would usually be considered a 'phonemic' difference from /i:/ to /ɪ/ to influence their interpretation of the morphological structure of a sentence spoken by a particular talker.

Accent listeners also partially generalised their training from prefixes to stems. As discussed in Section 4.5.2, the mechanisms that enable the co-existence of specificity and generalisation are unclear. They may include a task-related strategy (e.g., a 'tolerant' approach) interacting with morphologically-sensitive perceptual learning. Alternatively, listeners may learn associations of varying strengths at various levels of linguistic structure concurrently. For example, Accent subjects may have learned a strong association with the prefix *re-* and a weaker association with any syllable /ri:/, irrespective of its other properties.

The latter explanation relates closely to Properties 3 and 7, and is consistent with the results of a recent study by Dahan and Mead (2010). They demonstrated that the extent of listeners' adaptation to noise-vocoded speech relates to the similarity of the training and test situations in several respects. Listeners were trained to understand noise-vocoded speech by hearing monosyllabic stimuli in vocoded and unaltered forms. Learning generalised to novel stimuli, but was modulated by the degree to which novel stimuli resembled training stimuli in terms of phoneme identity, position

within the syllable, the adjacent phonetic context and (less conclusively) the talker. Therefore, listeners were able to learn about the effects of vocoding at many levels of structure concurrently.

Further support for perceptual learning at multiple levels of representation is provided by Jesse et al. (2007). Their results showed that talker-familiarity decreased reaction times in a lexical decision task for repeated words and also, but to a lesser extent, words containing repeated phonemes. This implies some perceptual learning of talker-specific detail at a phonemic level and a stronger learning effect at another level of representation – possibly the word level, or perhaps learning allophones relating to a position in syllable structure, or learning of idiosyncratic forms of coarticulation.

Concurrent perceptual learning of associations of differing strengths at many levels of structure depends on listeners' ability to learn probabilistic associations (as opposed to categorical associations, in which every instance of the relevant context exemplifies the atypical pronunciation). For example, in Experiment 3, Accent listeners were exposed to all *re-* prefixes pronounced as [rɪ] during training, which means that, by definition, they were also exposed to a proportion of /ri:/ syllables pronounced as [rɪ]. That is, they were exposed to a probabilistic association between the [rɪ] pronunciation and /ri:/ syllables. If the consistency of an association affects learning (see Property 7), it is unsurprising that listeners should have learned a strong association of the [rɪ] pronunciation with the prefix *re-*, and a weaker association of the [rɪ] pronunciation with the syllable /ri:/. Taking this logic further, because Accent subjects heard /ri:/ as [rɪ] in prefixes, they were exposed to a small subset of /i:/ phonemes pronounced as [ɪ], and so there is likely to have been an even weaker association of /i:/ with [ɪ]. Further research is necessary to test this possibility; for example, investigating whether listeners' performance on 'mispronunciations' in different contexts is affected by the training (see Section 4.5.2.1 for examples).

Many investigations of perceptual learning have involved subjects adapting to a phonetic characteristic that is absolutely categorical. However, there is some evidence that listeners can learn probabilistic phonetic associations (Docherty et al., 2008; Goldrick, 2004; Kraljic et al., 2008b). Experiment 2a provides further support. Listeners adapted to the systematic association between duration and prosodic phrase

position even though the confound between grammatical category and prosodic phrase position produced a probabilistic rather than an absolutely categorical association with duration – nouns always occurred in phrase-final position, but verbs were sometimes phrase-final and sometimes not.

When acquiring a language, listeners will be exposed to a range (whether large or small) of accents, to different talkers, to slips of the tongue and to natural variation across tokens. In this context, it is unsurprising that listeners have the ability to adapt to probabilistic associations, and some kind of probabilistic approach can be considered necessary to any theory of speech perception. (The reasons why, in Experiment 2a, subjects learned the probabilistic rather than the categorical association are discussed in Section 5.4.2.2.)

The embedding of smaller units within larger ones, or of more specific contextual associations within broader ones, is pervasive in speech. For example, /ɔ:/ is (or can be) a phone, a syllable and a word; lengthening at the end of intonational phrases also implies that a proportion of accentual phrases, words and syllables exhibit final lengthening; furthermore, talkers are 'embedded' within systems of social class, gender and communities. It seems probable that this embedded structure, coupled with listeners' ability to learn probabilistic associations, leads to perceptual learning of different strength associations at many different levels of representation.

There is much scope for research into perceptual learning of phonetic associations with multiple levels of representation, both linguistic and non-linguistic. Perceptual learning has yet to be demonstrated in relation to many contexts, such as the way a talker signals the end of a speech turn or their emotional state, or pragmatic aspects of a situation such as whether a talker is drunk or ill, or social aspects of communication such as gender or class. To test the hypothesis of concurrent perceptual learning at multiple levels of non-linguistic structure, one could investigate whether (and under what circumstances) perceptual learning relating to a talker generalised to other talkers that were similar in certain ways, e.g. gender, age, social class or group affiliation. Such learning effects may well differ according which aspects of a social situation are particularly relevant to the listener at that time (see Section 5.4.2.3).

5.4 FACTORS INFLUENCING THE LEARNABILITY OF AN ASSOCIATION

Property 4 proposed that the extent to which a novel association will be learned is dependent on the degree of activation of the phonetic properties and the degree of activation of the relevant categories. The following sections discuss how evidence from the present experiments and previous studies sheds light on the factors that influence activation levels and thus learnability. Section 5.4.1 deals with factors affecting the activation of phonetic properties (relating to Property 5), while Section 5.4.2 discusses factors affecting the activation of previously-learned categories or structures (relating to Property 6). The theoretical implications of these influences are discussed in Sections 5.4.3–4.

5.4.1 Activation of phonetic properties

Obviously, to be learnable, a systematic phonetic pattern must be perceptible. However, within that constraint, there are many degrees of perceptual salience, and many factors that influence perceptual salience. Property 5 proposed that the degree of activation of a phonetic property is determined by its perceptual salience in the current task, and that this influences the degree of perceptual learning. Experiment 1 provides some evidence that greater perceptual salience can lead to perceptual learning of stronger associations.

The results of Experiment 1 showed that syllable stress has a strong influence on the degree of perceptual learning of an atypical variant: hearing an atypical pronunciation in stressed syllables had a stronger perceptual learning effect than hearing it in unstressed syllables. Syllable stress is known to relate to the perceptual salience of an atypical pronunciation. For example, Cole and Jakimik (1980) found that American subjects detected ‘mispronunciations’ almost twice as frequently when they occurred in stressed syllables. As discussed in Chapter 2 (Section 2.4.3.2), stressed syllables are more acoustically prominent and auditorily salient than unstressed syllables, and they are likely to carry a heavier information load, both of which will increase their perceptual salience relative to unstressed syllables, and make listeners more likely to adjust their phonetic categories.

It may seem self-evident that the perceptual salience of a phonetic property influences the degree of perceptual learning, but it has important implications for the extent of

perceptual learning in natural communication. In natural speech, the magnitude of a phonetic change is likely to be affected by its lexical stress and its position within the prosodic structure (e.g., whether it is in focus, accented, new vs. given information, function vs. content words), and this may affect perceptual salience and thus influence the degree of perceptual learning. Furthermore, conversational speech can occur in a variety of situations with varying levels of background noise that may affect the auditory salience of particular phonetic properties, and so affect the extent to which listeners are able make use of different kinds of phonetic and contextual cues (Mattys et al., 2005). In addition, as discussed in Section 2.1.1.3, the use of natural tasks during training in the present experiments was motivated by the important role that feedback and task demands can play by directing the listener's attention to particular aspects of the stimuli and increasing their perceptual salience. An important role for perceptual salience implies that studies investigating only isolated words or syllables in clear listening conditions may not provide a realistic indication of what is learnable from conversational speech.

Acknowledging the importance of perceptual salience in determining what is learned also highlights the role of native language background and accent background in perceptual learning. Prior knowledge is known to play an important role in perception. Babies as young as 6 months display language-specific patterns of discrimination that indicate differences in their perception of similarity (Kuhl et al., 1992), and best exemplar locations of vowels are known to vary with regional accent in adults (Evans & Iverson, 2004). The conclusion must be that the perceptual salience of an acoustic property, and hence the likelihood that it will be learned, will be influenced by the native language and accent of the listener.

5.4.2 Activation of prior categories

Experiments 2 and 3 provided some evidence that it is not only the activation level of phonetic properties that affects perceptual learning; the degree of activation of prior categories will also influence the extent to which they become associated with the input, as described in Property 6.

5.4.2.1 Category 'robustness'

The results of Experiment 3 support the proposal that the 'robustness' of a category influences the strength of a novel association. 'Robustness' relates to the number of attributes that co-occur to form a category, and the consistency with which they do so. The performance of Control subjects in Experiment 3 suggests that the semantic consistency of the morphological conditioning context may have influenced perceptual learning. As discussed in Section 4.5.3, Control subjects showed an intelligibility gain for atypically-realised prefixes but not for atypically-realised stems between the first and second half of the intelligibility-in-noise task. The consistent meaning of the *re-* prefix may have made it easier for listeners to learn the systematic phonetic pattern, while, in contrast, *re-* in stems does not have an independent meaning and thus it may have been harder for listeners to 'latch on' to the systematicity in the input. This interpretation is consistent with previous evidence for the importance of semantic correlates in learning. For example, in an artificial language learning experiment, Braine (1987) found that semantic correlates in a subset of each word class facilitated learning of word classes.

Semantic consistency is certainly not *necessary* for perceptual learning – many of the perceptual learning studies reviewed in Chapter 1 demonstrate learning of phonetic patterns that relate to the realisation of features or phonemes rather than to consistent semantic properties, so these studies relied on other properties (principally phonetic and phonological) that lead to the formation of coherent categories. However, other things being equal, association with a consistent meaning should enhance the probability of perceptual learning because it provides an additional 'dimension' of similarity which serves to link the input exemplar more strongly with similar exemplars that share this meaning than with exemplars that do not share the meaning.

It is likely that any aspects of prior experience that support the existence of a robust category will enhance the probability that associations with that category will be formed because more robust categories will be more easily activated. Semantic consistency is one factor that can increase the robustness of a category, but phonetic, phonological and visual consistency are alternative factors that are likely to contribute to category robustness (see e.g., Brooks et al., 1993, for the facilitatory effect of phonological consistency on learning of word classes).

Dupoux et al. (2008) show that the lack of a robust category representation can limit perceptual learning. In their study, French learners of Spanish had difficulty perceiving Spanish stress contrasts when asked to discriminate word–non-word minimal pairs contrasting only in stress (e.g., ˈ*gorro* ('hat') vs. *go*ˈ*rro*) in a lexical decision task. Good performance in an AX discrimination task with minimal pairs differing in stress position indicated that French listeners were able to perceive the acoustic differences under some circumstances, so the difficulty was not a lack of perceptual salience. The authors attribute French listeners' difficulties to a processing 'deafness' at a phonological level, due to the lack of a contrastive stress in French. In terms of Property 6, this could be described as perceptual learning being prevented by a lack of activation of a robust phonological category representation.

5.4.2.2 Bottom-up activation of categories because of prior associations

The results of Experiment 2 suggests that prior associations are another factor that influences the degree of activation of previously-learned higher-level categories or structures.

Experiment 2 showed that that listeners adapted to a probabilistic association between phrase-finality and duration, even though an alternative, absolutely categorical association between grammatical category (noun or verb) and duration was available in the speech signal. Duration is known to be a strong cue to prosodic phrase boundaries, so, as discussed in Chapter 3, the existence of this prior association between duration and prosodic structure may have increased the probability that listeners would learn the novel association between duration and phrase-finality.

This effect of a prior knowledge can be described in terms of the influence of bottom-up activation due to relatively strong prior associations (Property 6). Listeners' expectations regarding the information that duration conveys about prosodic structure may have increased the extent to which bottom-up durational cues activated prosodic units. The resulting higher activation of prosodic units allowed the probabilistic relationship between duration and phrase-finality in this experiment to override the categorical relationship between duration and grammatical category.

5.4.2.3 Other factors affecting activation of previously-learned representations

The activation of previously-learned representations is also likely to be affected by task and attention. Goldinger and Azuma (2003) found that unconscious and inexplicitly conveyed experimenter-bias towards a phoneme- or syllable-based view of perceptual units affected reaction times in phoneme- and syllable-monitoring, presumably because subjects focussed on different levels of representation during the task. This kind of focus on a particular level of representation may have facilitated perceptual learning for Control subjects in Experiment 3 – the intelligibility-in-noise task involved trying to understand the meaning of the sentences, and this could have emphasised semantic attributes of the stimuli relative to phonological attributes, making the 'prefix' category particularly highly activated.

Further research is required to investigate the effect on perceptual learning of drawing subjects' attention to (or away from) a particular level of representation.

5.4.3 How do the factors that affect learnability interact?

The properties outlined in Section 5.2 cannot always be used to predict (or explain) why listeners should learn one pattern and not another. For example, although listeners in Experiment 2a learned an association between the durational cue and prosodic structure, it is not possible determine whether NLVS or NSVL subjects (or both) adapted their responses to Long stimuli. Either explanation (or both) could be consistent with the properties outlined in Section 5.2. NLVS subjects may have learned the durational cue because it is an extreme form of a prior association between longer duration and phrase-finality, and therefore the bottom-up stimulation from the durational cue would have activated 'phrase-finality' strongly (see Section 5.4.2.2 and Property 6). Alternatively, the more rhythmically-unusual training for NSVL subjects may have directed their attention to the rhythmic properties of the speech, thus enhancing perceptual learning because of increased activation of the durational cue (see Property 5). There is evidence that listeners can learn about phonetic characteristics that are reasonably consistent with prior experience (e.g., Sjerps & McQueen, 2010; Docherty et al., 2008), but also that listeners can learn about relatively 'odd' characteristics (e.g., Norris et al., 2003; Best, 1995). An additional control condition with no duration manipulations during training would have helped resolve this question in Experiment 2.

This example illustrates that it is not sufficient to understand that factors such as prior associations and directed-attention can influence what is learned; it is also necessary to understand how they interact. It is only when the combined effects of prior knowledge and current task demands have been established that it will be possible to predict what listeners will learn from exposure to a particular stimulus in a particular situation.

5.4.4 The stability-plasticity relationship: the role of prior knowledge

Understanding the balance between plasticity and stability is crucial to building a realistic model of speech perception. A polysystemic exemplar-based approach to speech perception would potentially allow any type of audible phonetic variation to become associated with any context. However, to prevent over-generalisation and maintain a relatively stable perceptual system, there must be mechanisms governing how novel information is integrated with prior knowledge. The present experiments begin to shed light on the role of prior knowledge in constraining or enhancing perceptual learning, and so help to elucidate the plasticity-stability relationship in speech perception.

Prior knowledge seems to affect perceptual learning in several ways. Firstly, it affects the perceptual salience of phonetic properties, as illustrated by Experiment 1 (see Section 5.4.1). Secondly, as suggested by the results of Experiment 2, novel associations between a phonetic cue and a category may be more likely to be learned in cases where prior experience has shown that the cue is informative about that category, and, conversely, long-term associations are unlikely to be overturned unless the evidence is very strong (see Section 5.4.2.2). Finally, the robustness of previously-learned categories is likely to have a role in perceptual learning, as discussed in relation to the role of semantic consistency in facilitating Control subjects' learning of the atypical pronunciation in prefixes in Experiment 3 (see Section 5.4.2.1).

The mechanisms governing how novel information is integrated with prior knowledge should be a key area of future research, and the properties outlined in Section 5.2 could provide a framework for investigating the effects of prior knowledge on perceptual learning in terms of its effects on the activation of phonetic properties (Property 5) and on the activation of previously-learned categories (Property 6).

Cross-linguistic and cross-dialectal studies could elucidate the stability-plasticity relationship by providing a means of manipulating both the perceptual salience of certain phonetic cues, and also the robustness of certain previously-learned phonological categories.

Studies of child language acquisition are also potentially informative about the role of prior knowledge in influencing what is learned. For example, White et al. (2008) showed different learning outcomes from exposing children of different ages to the same stimuli (see Section 1.4.3), and Vihman and colleagues (e.g. Vihman & Croft, 2007) have suggested that an individual child's 'phonological templates' can lead to different developmental trajectories.

5.5 METHODOLOGICAL IMPLICATIONS OF THE PRESENT RESEARCH

In order to maximise the relevance of the results to everyday speech perception, the experiments in this thesis used relatively natural stimuli and tasks. The results highlight both practical and theoretical considerations that should be taken into account when designing future perceptual learning experiments, particularly within a naturalistic framework.

5.5.1 Multiple levels of representation

Acknowledging the importance of multiple levels of structure in perceptual learning has implications for the design and interpretation of future studies. Many investigations into what listeners learn from exposure to a stimulus seem, at least implicitly, to be searching for a single locus of perceptual learning. For example, Eisner and McQueen (2005, p. 224) claim that 'perceptual learning about idiosyncratic speech is applied at a segmental level and is, under these exposure conditions, talker specific'. Likewise, Kraljic and Samuel (2006, p. 267) argue that 'listeners are able to apply learning from one phoneme and speaker to new phonemes and new speakers, indicating that perceptual learning occurs at the featural level'. The present results, along with the studies discussed in Section 5.3, suggest that listeners may learn multiple new associations at many different levels of structure at the same (or much the same) time, and therefore an approach is needed that not only

acknowledges the multiplicity of category membership of any given chunk of sound, but also explicitly investigates learning in relation to multiple levels of representation.

Systematic investigation of what listeners have learned at multiple levels of representation was admirably demonstrated by Dahan and Mead (2010; see Section 5.3), but has been somewhat lacking in previous studies. For example, as noted in Section 1.4.6, Kraljic et al. (2008b) claimed that listeners failed to learn about a phonetic characteristic (an ambiguous [s/ʃ]) that was conditioned by the talker having a pen in their mouth. However, their assessment of perceptual learning involved categorisation on an auditory-only continuum, and therefore it would not have been clear to listeners whether the talker had a pen in their mouth or not in this task. On the assumption that listeners use information only when it is relevant, it would be necessary to include sight of the talker chewing on a pen as a conditioning factor in the assessment task, in order to test whether a 'pen-related' association had been learned.

Likewise, Maye et al.'s (2008) study of adaptation to vowel lowering (described in Section 1.4.2) could be extended to test whether listeners learn about particular vowels or a more general category of front vowels, or both. And, as discussed in Section 4.5.2.1, to elucidate precisely what was learned in Experiment 3, it would be necessary to include the atypical variant in more phonetic and linguistic contexts in the intelligibility-in-noise task.

5.5.2 Generalisability of conclusions

It is important to establish what perceptual learning studies seek to achieve. Is the aim to determine what is ultimately learnable by the listener, or is it to find out the kind of characteristics that listeners are likely to adapt to in everyday conversation? Both questions are interesting, but they require different research methods, and researchers should be explicit about the applicability of their work to natural speech.

Section 5.2 proposed that what is learned by listeners from exposure to a stimulus depends on activation levels that are influenced by the listeners' prior experience, by his or her attention, and by the auditory salience of aspects of the speech signal. This multiplicity of influences highlights the limitations of studies that use unnatural stimuli or tasks to address perceptual learning in ordinary listening situations – they

only show what can be learned (or not) by a particular group of listeners exposed to specific stimuli in a certain situation. The effects of task on attention have been established (e.g., Loebach et al., 2008, see Section 2.1.1.3; Nosofsky, 1986), and it is becoming clear that remarkably subtle aspects of an experimental situation can affect listeners' interpretation of stimuli. Goldinger and Azuma's (2003) demonstration of unconscious and indirect experimenter bias has been discussed above, and Hay and Drager (2010) provide another compelling example. They required listeners to match a natural vowel /ɪ/ spoken by a New Zealander to a vowel on a synthetic vowel continuum going from raised and fronted Australian-like tokens to lowered and centralised New-Zealand-like tokens. Perception of the vowels shifted depending on whether a stuffed toy koala (associated with Australia) or a stuffed toy kiwi (associated with New Zealand) was present on the desk during the experiment.

The present research has shown that it is challenging but still possible to study perceptual learning within a relatively natural context. The results of Experiments 1 and 3 show that listeners learned novel phonetic associations from a story spoken by a real talker, or, in Experiment 2, by a high-quality resynthesised natural voice. Although the present research did not use conversational speech, the training tasks in the present research aimed to incorporate the sources of top-down information that would be available to the ordinary listener in conversation. This was achieved (as far as possible) by exposing listeners to a story. Unlike, for example, Norris et al. (2003), there was no meta-linguistic training task, and, in contrast to many studies (e.g., Allen & Miller, 2004; Dahan & Mead, 2010; Francis et al., 2000), there was no explicit feedback. These experiments are therefore more informative than many previous studies about the kind of associations that listeners may learn in ordinary conversation.

Nevertheless, the limits in generalisability of the present research are acknowledged: conclusions about what is learned apply to young British English listeners paying close attention to stories that are spoken in clear-speech SSBE (or mildly Leeds-accented English in Experiment 1) that has slightly atypical characteristics. More work is required in order to determine whether the listeners will learn the same associations in the context of, for example, background noise, speaker-listener

interaction or divided attention, and whether listeners from other language or dialect backgrounds will behave in the same way.

Likewise, negative results should not be uncritically generalised to other learning contexts – that an association is not learned in one experimental task does not necessarily imply that it is unlearnable. It may be learnable by other listeners or in other contexts. For example, it may be possible for listeners to learn an association between nouns and verbs and a durational cue, even though listeners did not do this in Experiment 2. For instance, if the training comprised word lists rather than sentences then the confound between prosodic structure and syntactic class would be eliminated, and learning a phonetic association with syntactic class might be more probable because listeners would not have alternative, 'more plausible' associations to learn.

Until a more complete understanding is achieved of how prior knowledge and task demands affect perceptual learning, researchers should design their experiments to investigate the specific situations in which they are interested, and should be cautious about generalising their specific results to different situations.

5.5.3 Minimising or monitoring perceptual learning during assessment

Assessment tasks are intended to assess what has been learned during training, but implicit training within the assessment task risks obscuring between-group differences. Experiments 1 and 3 highlight how difficult it can be to design an assessment task that minimises perceptual learning.

Experiment 1 underlines the power of lexical status as an influence on adaptation, perhaps especially within a lexical decision task: Control subjects' knowledge that the talker did not exhibit the atypical variant in the training story did not prevent them from adapting to the atypical variant in the test task, guided by lexical information.

In Experiment 3, the intelligibility-in-noise task was intended to limit perceptual learning by reducing the reliability of information available from higher levels of linguistic structure. However, the performance of Control subjects indicates that it was not wholly successful – Control subjects adapted to the atypical pronunciation in prefixes between the first and second halves of the experiment. Thus, even degraded feedback from higher levels of structure can guide perceptual learning.

Experiment 2 was more successful at minimising the use of top-down expectations to guide learning during assessment. Experiments 2a and 2b used sentence completion tasks in which the interpretation of the critical phonetic characteristic was dependent on the listener, and so there was no 'correct' or 'incorrect' response that could direct learning.

It is not always possible to use listener interpretation to avoid the problem of implicit learning during assessment (cf. Experiment 3). In such cases, the present experiments demonstrate the usefulness of monitoring subjects' performance over the course of the experiment.

5.5.4 How perceptual learning effects are manifested

The present research raises interesting questions about how the effects of perceptual learning are manifested in listeners' behaviour. Many previous studies of perceptual learning, and Experiment 1 in this thesis, have used assessment tasks that require listeners to make explicit decisions about phonological categories (e.g., Norris et al., 2003; Kraljic & Samuel, 2005, 2006), lexical status (Maye et al., 2008) or social groups (Docherty et al., 2008). Although Experiments 2 and 3 assessed perceptual learning in a more natural way by trying to access listeners' interpretation of the meaning of the stimuli, they nevertheless followed the general approach of assessing perceptual learning by monitoring listeners' linguistic categorisation of certain parts of the stimulus, whether as a noun or verb or as a prefix or part of a stem.

Experiments 1 and 3 suggest that perceptual learning can be manifested in other ways than by changing listeners' linguistic category judgements. Experiment 1 provided tentative evidence that perceptual learning can affect the percept of talker similarity (see Section 2.4.3.1), while Experiment 3 suggests that there was a subtle effect of perceptual learning of a prefix-related phonetic characteristic on the intelligibility of the suffix, presumably due to more efficient processing of the prefix by Accent subjects (see Section 4.5.1.1).

This suggests that a complete picture of the effects of talker-familiarity can only be achieved by using assessment tasks that go beyond linguistic category judgements relating to the immediate environment of the atypical variant. For example, assessments of accent similarity may reveal more subtle learning effects.

5.6 CONCLUSION

This research has demonstrated that listeners are sensitive to the prosodic and morphological context in which atypical pronunciations occur, and they are able to learn novel associations between phonetic detail and higher levels of linguistic structure. This is consistent with a polysystemic exemplar-based approach to speech perception, in which the speech signal is processed with reference to many levels of representation concurrently. Evidence that listeners can adapt to 'context-sensitive' variants helps to escape from a conception of exemplar models in which the effect of recent experience on perception is a relatively automatic consequence of auditory processing. Instead, it emphasises the extent to which what is learned relies on processing of the entire communicative situation in the context of the listener's past experience.

The present experiments have also shown that certain patterns of phonetic detail are more learnable than others. The proposed properties of a model of perceptual learning (Section 5.2) suggest how factors such as attention, prior knowledge and auditory salience can influence what is learned by affecting the perceptual salience of phonetic properties and the degree of activation of previously-learned categories.

There remains much scope for research into the phonetic associations that listeners are able to learn. Specific possibilities for future research have been mentioned as they arose in the discussion. Perhaps the most exciting area for future research is the role of prior knowledge in influencing what is learned, because this is what governs the relationship between plasticity and stability in the perceptual system.

BIBLIOGRAPHY

Adank, P., & McQueen, J. M. (2007). The effect of an unfamiliar regional accent on spoken word comprehension. In J. Trouvain & W. J. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1925–1928). University of Saarbrücken, Saarbrücken.

Allen, J. S., & Miller, J. L. (2004). Listener sensitivity to individual talker differences in voice-onset-time. *Journal of the Acoustical Society of America*, *115*, *6*, 3171–3183.

Alfonso, P. J., & Baer, T. (1982). Dynamics of vowel articulation. *Language and Speech*, *25*, *2*, 151–173.

Baayen, R. H. (2004). Statistics in psycholinguistics: A critique of some current gold standards. *Mental Lexicon Working Papers I*, University of Alberta, Edmonton, 1–45.

Baayen, R. H. (2007). *Analyzing Linguistic Data. A Practical Introduction to Statistics using R.* Cambridge: Cambridge University Press.

Baker, R. (2008). *The Production and Perception of Morphologically and Grammatically Conditioned Phonetic Detail*. Unpublished PhD dissertation, University of Cambridge.

Beach, C. M. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity: Evidence for cue trading relations. *Journal of Memory and Language*, *30*, 644–663.

Beddor, P. S., Krakow, R. A., & Lindemann, S. (2001). Patterns of perceptual compensation and their phonological consequences. In E. V. Hume, & K. Johnson (Eds.), *The Role of Speech Perception in Phonology* (pp. 55–78). San Diego: Academic Press.

Bertelson, P., Vroomen, J., & de Gelder, B. (2003). Visual recalibration of auditory speech identification: A McGurk aftereffect. *Psychological Science*, *14*, *6*, 592–597.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 171–204). Baltimore: York Press.

Boersma, P., & Weenink, D. (2007). Praat: doing phonetics by computer [Computer program]. Version 4.6.06, retrieved 20 June 2007 from http://www.praat.org/

Bond, Z. S. (1981). Listening to elliptic speech: Pay attention to stressed vowels. *Journal of Phonetics*, *9*, 89–96.

Bradlow, A. R., & Bent, T. (2007). Perceptual adaptation to non-native speech. *Cognition*, *106*, *2*, 707–729.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate, and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics*, *61*, *2*, 206–219.

Braine, M. D. S. (1987). What is learned in acquiring word classes: A step toward an acquisition theory. In B. MacWhinney (Ed.) *Mechanisms of Langauge Acquisition* (pp. 65–86). Hillsdale, NJ: Erlbaum.

Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R. E., & Sudhalter, V. (1993). Acquisition of gender-like noun subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language*, *32*, 76–95.

Bybee, J. L. (1996). The phonology of the lexicon: evidence from lexical diffusion. In M. Barlow, & S. Kemmer (Eds.), *Usage-based Models of Language* (pp. 65–85). Stanford: CSLI Publications.

Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, *30*, 348–369.

The Chambers Dictionary (1998). Edinburgh: Chambers Harrap Publisher Ltd.

Cheang, H. S., & Pell, M. D. (2008). The sound of sarcasm. *Speech Communication*, *50*, *5*, 366–381.

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: evidence from Korean. *Phonetica*, *58*, 129–162.

Cho, T., McQueen, J. M., & Cox, E. A. (2007). Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics*, *35*, 210–243.

Cho, T., & Keating, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, *29*, *2*, 155–190.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access 1. Adult data. *Journal of Memory and Language*, *51*, 523–547.

Clarke, C. M., & Garrett, M. F. (2004). Rapid adaptation to foreign-accented English. *Journal of the Acoustical Society of America*, *116*, *6*, 3647–3658.

Cole, R. A., & Jakimik, J. (1980). How are syllables used to recognize words? *Journal of the Acoustical Society of America*, *67*, *3*, 965–970.

Costello, J., & Onstine, J. (1976). The modification of multiple articulation errors based on distinctive feature theory. *Journal of Speech and Hearing Disorders*, *41*, 199–215.

Cooke, M. (2003). Glimpsing speech. *Journal of Phonetics*, *31*, *3–4*, 579–584.

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners. *Language and Speech*, *45*, *3*, 207–228.

Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge, MA: Harvard University Press.

Couper-Kuhlen, E., & Selting, M. (Eds.) (1996). *Prosody in conversation: interactional studies*. Cambridge: Cambridge University Press.

Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, *25*, *4*, 385–400.

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, *1*, 113–121.

Cutler, A., & Pearson, M. (1986). On the analysis of prosodic turn-taking cues. In C. Johns-Lewis (Ed.) *Intonation in Discourse* (pp. 139–155). London: Croom Helm.

Cutler, A., McQueen. J. M., Butterfield, S., & Norris, D. (2008). Pre-lexically driven perceptual retuning of phoneme boundaries. *Proceedings of Interspeech. Brisbane, Australia, 22–26 September*, 2056–2056.

Dahan, D., Drucker, S. J., & Scarborough, R. A. (2008). Talker adaptation in speech perception: Adjusting the signal or the representations? *Cognition*, *108*, 710–718.

Dahan, D., & Mead, R. L. (2010). Context-conditioned generalization in adaptation to distorted speech. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, *3*, 704–728.

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, *1*, 218–244.

Dell, G. S., Reed, K. D., Adams, D. R., & Meyer, A. S. (2000). Speech errors, phonotactic constraints and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26*, *6*, 1355–1367.

Deese, J., & Hulse, S. H. (1967). *The Psychology of Learning*. Third Edition. McGraw-Hill Inc.

Docherty, G. J. (2007). Speech in its natural habitat: Accounting for social factors in phonetic variability. In J. Cole, & J. I. Hualde (Eds.) *Laboratory Phonology 9* (pp. 1–35). Berlin: Mouton de Gruyter.

Docherty, G., Langstrof, C., & Foulkes, P. (2008). Perceptual evaluation of sociophonetic variability: how do listeners learn? *Poster presented at Laboratory Phonology 11*, *Wellington, New Zealand, 30 June – 2 July.*

Docherty, G., Foulkes, P., Tillotson, J., & Watt, D. (2006). On the scope of phonological learning: Issues arising from socially-structured variation. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.) *Laboratory Phonology 8* (pp. 393–421). Berlin: Mouton de Gruyter.

Dupoux, E., & Green, K. (1997). Perceptual adjustment to highly compressed speech: Effects of talker and rate changes. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, *3*, 914–927.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, *106*, 682–706.

Durieux, G., & Gillis, S. (2001). Predicting grammatical classes from phonological cues: An empirical test. In J. Weissenborn & B. Höhle (Eds.) *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, Vol. 1 (pp. 189–229). Amsterdam: John Benjamins.

Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception and Psychophysics*, *67*, *2*, 224–238.

Eisner, F., & McQueen, J. M. (2006). Perceptual learning in speech: Stability over time (L). *Journal of the Acoustical Society of America*, *119*, *4*, 1950–1953.

Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, *27*, 143–165.

Evans, B. G., & Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentences. *Journal of the Acoustical Society of America*, *115*, *1*, 352–361.

Fabricius, A. (2002). Weak vowels in modern RP: An acoustic study of happY-tensing and KIT/schwa shift. *Language Variation and Change*, *14*, 211–237.

Farmer, T. A., Christiansen, M. H., & Monaghan, P. (2006). Phonological typicality influences lexical processing. *Proceedings of the National Academy of Sciences*, *103*, 12203–12208.

Firth, J. R. (1948). Sounds and prosodies. *Transactions of the Philological Society*, 127–152.

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Reseach Methods, Instruments, and Computers*, *35*, *1*, 116–124.

Fougeron, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, *29*, 109–135.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, *101*, *6*, 3728–3740.

Foulkes, P. (2006). Phonological variation – A global perspective. In B. Aarts, & A. McMahon (Eds.), *Handbook of English Linguistics* (pp. 625–669). Oxford: Blackwell.

Foulkes, P., & Docherty, G. (2006). The social life of phonetics and phonology. *Journal of Phonetics*, *34*, *4*, 409–438.

Fowler, C. A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research, 46*, 127–139.

Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words for speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, *26*, 489–504.

Francis, A. L., Baldwin, K., & Nusbaum, H. C. (2000). Effects of training on attention to acoustic cues. *Perception and Psychophysics*, *62*, *8*, 1668–1680.

Frick, R. W. (1985). Communicating emotion: the role of prosodic features. *Psychological Bulletin*, *97*, *3*, 412–429.

Gerken, L. (1996). Prosody's role in language acquisition and adult parsing. *Journal of Psycholinguistic Research*, *25*, *2*, 345–356.

Giegerich, H. J. (1992). English phonology: An introduction. Cambridge: Cambridge University Press.

Goldinger, S. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, *5*, 1166–1183.

Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, *2*, 251–279.

Goldinger, S. (2007). A complementary-systems approach to abstract and episodic speech perception. In J. Trouvain, & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 49–54). University of Saarbrücken, Saarbrücken.

Goldinger, S., & Azuma, T. (2003). Puzzle-solving science: The quixotic quest for units in speech perception. *Journal of Phonetics*, *31*, 305–320.

Goldrick, M. (2004). Phonological features and phontactic constraints in speech production. *Journal of Memory and Language*, *51*, 586–603.

Gordon-Salant, S., & Fitzgibbons, P. J. (2001). Sources of age-related recognition difficulty for time-compressed speech. *Journal of Speech, Language, and Hearing Research*, *44*, 709–719.

Hawkins, S. (2010). Phonetic variation as communicative system: Perception of the particular and the abstract. In C. Fougeron, B. Kühnert, M. d'Imperio, & N. Vallée, (Eds.) *Laboratory Phonology 10: Variation, Phonetic Detail and Phonological Modelling* (pp. 479–510). Berlin: Mouton de Gruyter.

Hawkins, S., & Smith, R. (2001). Polysp: a polysystemic, phonetically-rich approach to speech understanding. *Italian Journal of Linguistics – Rivista di Linguistica*, *13*, 99–188.

Hay, J. (2003). *Causes and Consequences of Word Structure*. New York: Routledge.

Hay, J., & Bresnan, J. (2006). Spoken syntax: The phonetics of giving a hand in New Zealand English. *Linguistic Review*, *23*, *3*, 321–349.

Hay, J., Nolan, A., & Drager, K. (2006a). From *fush* to *feesh*: Exemplar priming in speech perception. *Linguistic Review*, *23*, *3*, 351–379.

Hay, J., Warren, P., & Drager, K. (2 006b). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*, *4*, 458–484.

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*, *4*, 865–892.

Hazan, V., Sennema, A., Iba, M., & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication*, *47*, 360–378.

Hebb, D. O. (1949). *The Organization of Behavior: A Neurophysiological Theory*. New York/London: Wiley.

Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, *92*, 67–99.

Hudson, T., McDougall, K., & Nolan, F. (2010). Voice similarity and the telephone. *Poster presented at BAAP Colloquium, London, 29–31 March*.

Hunnicutt, S. (1985). Intelligibility versus redundancy – Conditions of dependency. *Language and Speech*, 28, *1*, 47–56.

Jesse, A., McQueen, J. M., & Page, M. (2007). The locus of talker-specific effects in spoken-word recognition. In J. Trouvain, & W. J. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1921–1924). University of Saarbrücken, Saarbrücken.

Johnson, K. (1997). Speech perception without speaker normalization. In K. Johnson, & J. Mullennix (Eds.) *Talker variability in Speech Processing* (pp. 145–165). San Diego: Academic Press.

Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, *34*, *4*, 485–499.

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics*, *24*, *4*, 359–384.

Jones, D. (1997). *English Pronouncing Dictionary*. 15th edition. (Ed.) Roach, P. and Hartman, J. Cambridge: Cambridge University Press.

Jurafsky, D., Bell, A., Gregory, M., & Raymond. W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. In J. Bybee, & P. Hopper (Eds.) *Frequency and the Emergence of Linguistic Structure* (pp. 229–254). Amsterdam: John Benjamins.

Jurafsky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In S. Stevenson, & P. Merlo (Eds.) *Papers in Laboratory Phonology VII* (pp. 3–34). Berlin/New York: Mouton de Gruyter.

Jusczyk, P. W., Hohne, E. A., & Baumann, A. (1999). Infants' sensitivity to allophonic cues for word segmentation. *Perception & Psychophysics, 61*, 1465–1476.

Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoustical Science and Technology*, *27*, *6*, 349–353.

Kellman, P. J., & Garrigan, P. (2009). Perceptual learning and human expertise. *Physics of Life Reviews*, *6*, 53–84.

Kelly, M. H. (1988). Phonological biases in grammatical category shifts. *Journal of Memory and Language*, *27*, 343–358.

Kelly, M. H. (2004). Word onset patterns and lexical stress in English. *Journal of Memory and Language*, *50*, 231–244.

Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, *3*, 389–403.

Kemps, R.. Ernestus, M., Schreuder, R., & Baayen, R. H. (2005a). Prosodic cues for morphological complexity: The case of Dutch plural nouns. *Memory and Cognition*, *33*, 430–446.

Kemps, R., Wurm, L. H., Ernestus, M., Schreuder, R., & Baayen, R. H. (2005b). Prosodic cues for morphological complexity in Dutch and English. *Language and Cognitive Processes*, *20*, *1/2*, 43–75.

Kirk, C. (2000). Syllabic cues to word segmentation. In A. Cutler, J. M. McQueen, & R. Zondervan (Eds.) *Proceedings of the Workshop on Spoken Word Access Processes (SWAP)* (pp. 131–134). Nijmegen, Max-Planck Institute for Psycholinguistics.

Klatt, D. (1975). Vowel lengthening is syntactically determined in a connected discourse. *Journal of Phonetics*, *3*, 129–140.

Klatt, D. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, *59*, *5*, 1208–1221.

Kraljic, T., & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*, 141–178.

Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning of speech. *Psychonomic Bulletin and Review*, *13*, *2*, 262–268.

Kraljic, T., Brennan, S. E., & Samuel, A. G. (2008a). Accommodating variation: Dialects, idiolects, and speech processing. *Cognition*, *107*, 54–81.

Kraljic, T., Samuel, A. G., & Brennan, S. E. (2008b). First impressions and last resorts: How listeners adjust to speaker variability. *Psychological Science*, *19*, *4*, 332–338.

Kuhl, P. K., Williams, K. A., Lacerda, F, Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*, 606–608.

Lachs, L., McMichael, K., & Pisoni, D. B. (2003). Speech perception and implicit memory: evidence for detailed episodic encoding of phonetic events. In J. S. Bowers, & C. J. Marsolek (Eds.) *Rethinking Implicit Memory* (pp. 215–235). Oxford: Oxford University Press.

Laver, J. (1980). *Phonetic Description of Voice Quality*. Cambridge: Cambridge University Press.

Laver, J. (2003). Three semiotic layers of spoken communication. *Journal of Phonetics*, *31*, *3*, 413–415.

Lavoie, L. (2002). Some influences on the realisation of *for* and *four* in American English. *Journal of the International Phonetic Association*, *32*, *2*, 175–202.

Lehiste, I., Olive, J. P., & Streeter, L. A. (1976). Role of duration in disambiguating syntactically ambiguous sentences. *Journal of the Acoustical Society of America*, *60*, 1199–1202.

Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, *6*, 172–187.

Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and lexical segmentation in hypokinetic and ataxic dysarthria. *Journal of the Acoustical Society of America*, *112*, *6*, 3022–3030.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, *94*, *3*, 1242–1255.

Local, J. (2003). Variable domains and variable relevance: interpreting phonetic exponents. *Journal of Phonetics*, *31*, *3*, 321–339.

Local, J. (2007). Phonetic detail and the organisation of talk-in-interaction. In J. Trouvain, & W. J. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1–10). University of Saarbrücken, Saarbrücken.

Loebach, J. L., Bent, T., & Pisoni, D. B. (2008). Multiple routes to the perceptual learning of speech. *Journal of the Acoustical Society of America*, *124*, *1*, 552–561.

Logan, G. D., Taylor, S. E., & Etherton, J. L. (1996). Attention in the acquisition and expression of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *22*, *3*, 620–638.

Mann, V. A., & Repp, B. H. (1980). Influence of vocalic context on perception of the /ʃ/-/s/ distinction. *Perception and Psychophysics*, *28*, *3*, 213–228.

Manuel, S., Y. (1995). Speakers nasalise /ð/ after /n/, but still hear /ð/. *Journal of Phonetics*, *23*, 453–476.

Marslen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*, *1*, 29–63.

Martin, J. G. (1970). On judging pauses in spontaneous speech. *Journal of Verbal Learning and Verbal Behavior*, *9*, 75–78.

Mattys, S. L., & Juscyzk, P. W. (2001). Do infants segment words or recurring contiguous patterns? *Journal of Experimental Psychology: Human Perception and Performance*, *27*, *3*, 644–655.

Mattys, S. L., White, L., & Melhorn, J. F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, *134*, *4*, 477–500.

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The Weckud Wetch of the Wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*, *3*, 543–562.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746–748.

McQueen, J. M., Cutler, A., & Norris, D. (2006). Phonological abstraction in the mental lexicon. *Cognitive Science*, *30*, 1113–1126.

McReynolds, L. V., & Bennett, S. (1972). Distinctive feature generalization in articulatory training. *Journal of Speech and Hearing Disorders*, *37*, 462–470.

Miller, J. L., & Liberman, A. M. (1979). Some effects of later occurring information on the perception of stop consonant and semivowel. *Perception and Psychophysics*, *25*, 457–465.

Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, *22*, *6*, 898–909.

Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the on-line syntactic analysis of spoken sentences. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*, 874–885

Morosan, D. E., & Jamieson, D. G. (1989). Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task. *Journal of Speech and Hearing Research*, *32*, 501–11.

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, *85*, 365–378.

Nagel, H. N., Shapiro, L. P., Tuller, B., & Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *Journal of Psycholinguistic Research*, *25*, *2*, 319–344.

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*, *1*, 62–85.

Norris, D. (1994). Shortlist: a connectionist model of speech recognition. *Cognition*, *52*, *3*, 189–234.

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*, *3*, 299–325.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology*, *47*, 204–238.

Norris, D. and McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, *2*, 357–395.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, *1*, 39–57.

Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory and Cognition*, *30*, *4*, 583–593.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception and Psychophysics*, *60*, *3*, 355–376.

Onishi, K. H., Chambers, K. E., & Fisher, C. (2002). Learning phonotactic constraints from brief auditory experience. *Cognition*, *83*, B13–B23.

Peperkamp, S., & Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. In J. Cole, & J. Hualde (Eds.), *Laboratory Phonology 9* (pp. 315–338). Berlin: Mouton de Gruyter.

Pickett, J. M. (1999). *The Acoustics of Speech Communication*. Needham Heights, MA: Allyn and Bacon.

Pierrehumbert, J. B. (1993). Prosody, intonation and speech technology. In M. Bates, & R. M. Weischedel (Eds.), *Challenges in natural language processing* (pp. 257–282). Cambridge: Cambridge University Press.

Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, *46*, *2–3*, 115–154.

Pierrehumbert, J. B. (2006). The next toolkit. *Journal of Phonetics*, *34*, *4*, 516–530.

Pierrehumbert, J. B., & Beckman, M. (1988). Japanese tone structure. *Linguistic Inquiry Monographs*, *15*, 1–282.

Pisoni, D. B. (1997). Perception of synthetic speech. In J. P. H. Van Santen, R. W. Sproat, J. P. Olive, & J. Hirschberg (Eds.) *Progress in speech synthesis* (pp. 541–556). New York: Springer-Verlag.

Plug, L. (2005). From words to actions: the phonetics of *eigenlijk* in two communicative contexts. *Phonetica*, *62*, 131–165.

Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of the Acoustical Society of America*, *90*, *6*, 2956–2970.

Pulvermüller, F. (2002). A brain perspective on language mechanisms: from discrete neuronal assembles to serial order. *Progress in Neurobiology*, *67*, 85–111.

Quené, H. (1993). Segment durations and accent as cues to word segmentation in Dutch. *Journal of Phonetics*, *94*, *4*, 2027–2035.

R Development Core Team (2007, 2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org

Remez, R. E., Fellowes, J. M., & Rubin, P. E. (1997). Talker identification based on phonetic information. *Journal of Experimental Psychology: Human Perception and Performance*, *23*, 651–666.

Saffran, J. R., & Thiessen, E. D. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*, *3*, 484–494.

Samuel, A. G. & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception, & Psychophysics, 71*, 1207–1218.

Schacter, D. L., & Church, B. A. (1992). Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, *5*, 915–930.

Scharinger, M., & Lahiri, A. (2010). Height differences in English dialects: Consequences for processing and representation. *Language and Speech*, *53*, *2*, 245–272.

Schepman, A., & Rodway, P. (2000). Prosody and parsing in coordination structures. *The Quarterly Journal of Experimental Psychology*, *53A*, *2*, 377–396.

Scott, D. (1982). Duration as a cue to the perception of phrase boundaries. *Journal of the Acoustical Society of America*, *71*, 966–1007.

Scott, D. R., & Cutler, A. (1984). Segmental phonology and the perception of syntactic structure. *Journal of Verbal Learning and Verbal Behavior*, *23*, 450–466.

Sebastián-Gallés, N., Dupoux, E., Costa, A., & Mehler, J. (2000). Adaptation to time-compressed speech: Phonological determinants. *Perception and Psychophysics*, *62*, 4, 834–842.

Sereno, J. A. (1986). Stress pattern differentiation of form class in English. *Journal of the Acoustical Society of America*, *79*, S36.

Sereno, J. A. (1994). Phonosyntactics. In L. Hinton, J. Nichols, & J. Ohala (Eds.) *Sound Symbolism* (pp. 263–275). Cambridge: Cambridge University Press.

Sereno, J. A., & Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research*, *19*, 387–404.

Sereno, J. A., & Jongman, A. (1995). Acoustic correlates of grammatical class. *Language and Speech*, *38*, *1*, 57–76.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, *25*, *2*, 193–247.

Shatzman, K. B., & McQueen, J. M. (2006). The modulation of lexical competition by segment duration. *Psychonomic Bulletin and Review*, *13*, *6*, 966–971.

Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*, B11–B21.

Shinn, P. C., Blumstein, S. E., & Jongman, A. (1985). Limitations of context-conditioned effects in the perception of [b] and [w]. *Perception and Psychophysics*, *38*, 397–407.

Sjerps, M. J., & McQueen, J. M. (2010). The bounds on flexibility in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, *1*, 195–211.

Smith, R. (2007). The effect of talker familiarity on word segmentation in noise. In J. Trouvain, & W. J. Barry (Eds.) *Proceedings of the 16th International Congress of Phonetic Sciences* (pp. 1917–1920). University of Saarbrücken, Saarbrücken.

Smith, R. (submitted). Production and perception of speaker-specific phonetic detail at word boundaries. Under review, *Journal of Phonetics*.

Smith, R., & Hawkins, S. (2000). Allophonic influences on word-spotting experiments. In A. Cutler, J. M. McQueen, & R. Zondervan (Eds.) *Proceedings of the Workshop on Spoken Word Access Processes (SWAP)* (pp. 139–142). Nijmegen, Max-Planck Institute for Psycholinguistics.

Sommers, M. S., Nygaard, L. C., & Pisoni, D. B. (1994). Stimulus variability and spoken word recognition. I. Effects of variability in speaking rate and overall amplitude. *Journal of the Acoustical Society of America*, *96*, 1314–1324.

Speer, S. R., Kjelgaard, M. M., & Dobroth, K. M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, *25*, *2*, 249–271.

Spinelli, E., Welby, P., & Schaegis, A.-L. (2007). Fine-grained access to targets and competitors in phonemically identical spoken sequences: the case of French elision. *Language and Cognitive Processes*, *22*, *6*, 828–859.

Stevens, K. N. (2002). Towards a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America*, *111*, *4*, 1872–1891.

Stevens, K. N., Manuel, S. Y., Shattuck-Hufnagel, S., & Liu, S. (1992). Implementation of a model for lexical access based on features. *Proceedings of the International Conference of Spoken Language Processing*, Banff, Alberta, Canada, 499–502.

Strange, W., & Dittman, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics*, *16*, *2*, 131–145.

Streeter, L. A. (1978). Acoustic determinants of phrase boundary location. *Journal of the Acoustical Society of America*, *64*, 1582–1592.

Tabossi, P., Collina, S., Mazzetti, M., & Zopello, M. (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 758–775.

Tagliapietra, L., Fanari, R., Collina, S., & Tabossi, P. (2009). Syllabic effects in Italian lexical access. *Journal of Psycholinguistic Research*, *38*, *6*, 511–526.

Turk, A.E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics, 27*, 171–206.

Van Donselaar, W., Koster, M., & Cutler, A. (2005). Exploring the role of lexical stress in lexical recognition. *The Quarterly Journal of Experimental Psychology A*, *58*, *2*, 251–273.

Vihman, M., & Croft, W. (2007). Phonological development: toward a "radical" templatic phonology. *Linguistics*, *45*, *4*, 683–725.

Walsh, T., & Parker, F. (1983). The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics*, *11*, 201–206.

Weil, S. A. (2001). Foreign-accented speech: Adaptation and generalization. *Unpublished MA Thesis*, Ohio State University.

Wells, J. (1982). Accents of English 3: Beyond the British Isles. Cambridge: Cambridge University Press.

Wells, J. (1999). British pronunciation preferences: A changing scene. *Journal of the International Phonetics Association*, *29*, 405–426.

West, P. (1999). Perception of distributed coarticulatory properties of English /l/ and /ɹ/. *Journal of Phonetics*, *27*, 405–426.

White, L., & Turk, A.E. (2010). English words on the Procrustean bed: Polysyllabic shortening reconsidered. *Journal of Phonetics, 38*, 459–471.

White, K. S., Peperkamp, S., Kirk, C., & Morgan, J. L. (2008). Rapid acquisition of phonological alternations by infants. *Cognition*, *107*, 238–265.

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, *91*, *3*, 1707–1717.

Wurm, L. H. (1997). Auditory processing of prefixed English words is both continuous and decompositional. *Journal of Memory and Language*, *37*, 438–461.

Zue, V. W. (1990). The use of speech knowledge in automatic speech recognition. In A. Waibel, & K.-F. Lee (Eds.) *Readings in Speech Recognition* (pp. 200–213). San Mateo: Morgan Kaufmann Publishing, Inc.

APPENDIX A

EXPERIMENT 1: TRAINING STORY

Key: **Bold underlined** = /ɪ/ in stressed syllable in stem

Bold = /ɪ/ in stressed syllable in affix

underlined italic = /ɪ/ in unstressed syllable in stem

italic = /ɪ/ in unstressed syllable in affix

Within the square brackets, the Global version is on the left, and the Stress-conditioned (and Control) on the right.

SECTION 1

Through the death of Dame **I**sabel Dav*i*s at the age of eighty-seven, we have lost a woman of great learn*i*ng, who has contr**i**but*e*d much to areas as diverse as natural h**i**story, eth*i*cs and anthropology. Even more [lamentable/d*i*stress*i*ng] may be the loss of a f**i**gure who was [truly/s*i*ncerely] liked by the [people/publ*i*c] of Br**i**tain. A [small/mod*e*st] funeral has been organised for friends and colleagues on Saturday the f**i**fteenth of March at a [peaceful/tranq*ui*l] location near the R**i**ver Thames. The funeral may well be followed by some form of national commemoration, but that has yet to be finalised. Rather than [a conventional statue/the *e*rection of a statue], a more appropriate tr**i**bute would perhaps be a nature conservation area financed by Dame **I**sabel's generous donation to the Conservation Trust.

Perhaps Dame **I**sabel was most well-known as a natural*i*st. Many botan*i*sts, birdwatchers and other [nature lovers/wildlife add*i*cts] look back fondly on her early broadcasts. These were among the first programmes ever to [be aimed at a young audience/have young people as the targ*e*t audience]. B**i**ll Oddie and Dav*i*d Bellamy are among those who have found Dame **I**sabel's 'Nature for the Young' [compell*i*ng/*i*nspir*i*ng], and now they too [aim/*i*ntend] to convey the [attraction/mag*i*c] and adventure of the natural world.

SECTION 2

Dame **I**sabel's own love of nature was [apparent/unm*i*stakeable] from an early age. As an **i**nfant, she qu**i**ckly learned to recognise the plants, birds and **i**nsects of the local area. Kather*i*ne Dav*i*s, **I**sabel's mother, once claimed that she knew her daughter [was unusually observant/showed unusual prom*i*se] when, at age four, she could accurately and cons**i**stently tell apart [moorhens/House Mart*i*ns] and [coots/Sand Mart*i*ns]. She also collect*e*d birds' eggs, an *i*llegal [hobby/pract*i*ce] nowadays, of course, but common among country boys and girls of the twenties.

But the young **I**sabel's zeal sometimes caused [problems/d*i*sruption]. On a 1970s chat show, M**i**ss*i*s Dav*i*s told of her daughter's unfortunate attempt to collect a rob*i*n egg from a tall sp**i**ndly tree on a v**i**ll*a*ge green some miles away. After a [frightened/pan*i*cked] search, **I**sabel was finally [found/d*i*scovered] unconscious

under the tree, after a branch had snapped on the way up and she had [unwisely/impulsively] tried to jump to the bottom. Fortunately, her [coat/pocket] had caught on a branch on the way down, [breaking/impeding] her fall. Her adventurous/intrepid spirit was unchanged, however, and the next year she had an equally close shave, after an ill-fated mission to find a rare [seaweed/puffin egg] caused [a fall/her to plummet] down a cliff, leading to three weeks of [painful treatment/anguish] and [several plaster casts/a vast number of bandages]!

SECTION 3

These early observations and adventures were the foundation of the [vital/essential] work she carried out as Chairman of the Conservation Trust of Britain, supporting various studies on, for instance, the spread of insect-borne [pathogens/infections] from Europe and the impact of modern farming [techniques/practices] on birds' nesting [routines/habits]. Dame Isabel was also first to [note/discover] the strange courtship practice of finches, where the birds raise the pitch of their calls by several tones while attracting a mate. Since then, the raised call has been [observed/noticed] across many species, from the chiff chaff to the [eagle/ostrich], and has called attention to the harm that noise pollution can cause to particular species. The Conservation Trust also led efforts to [find out/discover] [real/solid] proof of whether greater use of [Russian varieties of/hybrid] wheat was the cause of [falling/plummeting] skylark numbers, and what the actual outcome of the loss of arable field margins has been for populations of Red Poppies.

At a more hands-on level, she led the campaign against the construction of a [leisure centre/supermarket] on [Newbury/Trowbridge] meadows. The meadows are not [noteworthy/imposing] from an artistic viewpoint, but contain [large/immense] populations of [buttercups/water irises] over spring and a [good/splendid] [show/display] of [Meadowsweet/Marsh Orchids] over summer. The area was popular for Sunday afternoon strolls, and was well-used by the [locals/public]. Dame Isabel claimed that green spaces should be cherished not only for their [great/integral] beauty or interest, but also for the [vital/essential] fresh air and exercise that they can provide. The success of the campaign was [a considerable achievement/much to her credit], and Dame Isabel gave her support to several other projects. She had mixed success, but her commitment to the protection of green areas from unnecessary [harm/damage] never faltered.

SECTION 4

Dame Isabel claimed to have had a fairly normal childhood for her time. Nowadays we might consider that she had an advantaged upper middle class background; the household comprised a cook, housemaid and nanny, and both Isabel and her older brother, John, were sent away to [_/public] school. Unlike the [horror/horrid] stories of many sons and daughters sent away from their parents, the separation does not appear to have [upset/dismayed] them unduly. Isabel was particularly fond of biology lessons, but she was also good at [maths/physics], [German/Latin] and [French/classics]. Comments from friends [suggest/imply] that Isabel was happy at school, rarely incurring the wrath of the staff. "The only time she was ever [naughty/punished]," said [Martha/Alice] [Johnson/Bennett], "was when she found a stray [labrador/mastiff] and, unknown to her teachers, adopted the poor thing. She

made a home for Eric at the local parish church. She put a pile of old [clothes/blankets] under the [font/pulpit], and used the communion [cup/chalice] as a water bowl, but the dog made such a [noise/racket] that he was soon [found/discovered] by the local vicar. Fortunately, the staff were fairly [lenient/indulgent], but made clear that the diet of leftover soup and stale [bread/crumpets] that she'd been furtively smuggling to Eric were not sufficiently [wholesome/nourishing]. The school supplied a basket for Eric and gave the dog a temporary home under the [attic staircase/staffroom cupboard], while Isabel's punishment was to clean the church – a rather necessary job!"

SECTION 5

Dame Isabel's early love of nature was [obvious/noticeable] at school; she volunteered to dig the school pond and looked after the class [canary/rabbit]. She also campaigned vigorously for gardening lessons, and was then apparently inseparable from her vegetable patch, where she grew [superb/splendid] [broccoli/cabbage], [carrots/parsnips], and [leeks/radishes].

Her friends say she was an avid reader. "Nothing pleased her better," said [Martha/Alice] [Johnson/Bennett], "than to get away from the rigid constraints of everyday life by reading a good book. She was specially fond of The Phoenix and the Carpet, by E. Nesbit – we were the pre-Enid Blyton generation. Later, of course, she was a great fan of The Hobbit and the epic Lord of the Rings trilogy. I think her love of books fuelled her need for travel and adventure."

The school vacations were equally [action-packed/hectic]. The Davis household frequently travelled to Cumbria, where they stayed at a [small house/cottage] near Esthwaite Water. Thousands of sightseers must recognise the area, having visited Beatrix Potter's farmhouse nearby. She had bought the place after [writing/publishing] such [stories/classics] as [The Tailor of Gloucester/Peter Rabbit] and Squirrel Nutkin. Beatrix Potter and Dame Isabel were probably there at the same time for some periods, although Dame Isabel had no recollection of them meeting. Her memories were mainly of swimming around the cold waters of Lake Windermere and the unfair [lead due to/advantage of] her brother's extra three years as they raced each other to the [top/summit] of [Sca Fell Pike/Helvellyn]!

SECTION 6

But Dame Isabel's teenage years drew her attention away from the great outdoors. Her father, Philip Davis, was a [well-known/an illustrious] Tory MP of the thirties, and Isabel showed the classic teenage need to go against her elders by advocating a [complete/drastic] change of government. However, as we might [suppose/infer] from her usual idealistic fervour, she [eagerly/industriously] read a wide range of literature on the [subject/topic]. The Southampton Girls School magazine of 1935 contains [an assessment/a discussion] of John Stuart Mill's most important work, 'On Liberty'. Although idealistic, her essay contains the [frank/candid], sharp and [thoughtful/lucid] comments that were to emerge as the hallmark of her later work and help win her so many readers.

Isabel [was successful/flourished] at school, and went on to win a scholarship to St Hilda's College, Oxford. Throughout the late thirties, there were far fewer places

[allocated to/designated for] females than nowadays, and the award was [a noteworthy achievement/worthy of considerable merit]. However, shortly after the outbreak of World War Two, she insisted on turning down the award and joining the war effort, a move that caused a great deal of [annoyance/anguish] to her father. "My father was [furious/incensed]," she said. "He argued that anyone could sign up for the war effort, whereas very few had the necessary skills to study at the UK's most prestigious college, but I felt that, given [the current state of affairs/current events], there were more [useful/important] tasks that had to be [done/accomplished]."

SECTION 7

Her persistence paid off when her father realised that she was wholly intent on her purpose and he stopped [protesting/dissenting]. Dame Isabel spent the first two years of the war working at a munitions factory close to Cardiff**.** At the time, her location was kept from her parents and friends, due to the fact that the government wished to stop the factories from being targets for German bombers. But her stint at the factory [came to an abrupt end/finished abruptedly] when a small detonation caused [the loss of/severe damage] to the little finger of her right hand. Her injury meant that she was unable to do the complex finger operations that were [fundamental/basic] to the job. She left Wales to live at a friend's bedsit apartment, and spent the rest of the war working for the Met Office at their Head Office near Reading. Dame Isabel attributes her [superior/impressive] map-reading skills to her period at the Met. These were valuable throughout her later career, when she traversed terrain that was barely known to the Western world. "A quick eye and a good overall picture of the terrain was [vital/essential] to my later work," she said, "and constantly tracking the progress of weather fronts across the [ocean/Atlantic] gave me that instant vision that allowed me to [promptly/rapidly] pick out the crucial features on a map, and match them to the nearby landscape."

SECTION 8

The most painful aspect of Dame Isabel's war years was her concern for her brother John. John had been working as an [accountant/an insurance broker] at a London bank when he was conscripted. After eight weeks training he was sent to France, but [he served just a few months/he was part of just one minor skirmish] and was then was captured by the Nazis. There had been an [alarm/emergency] caused by a raid at the [troop's/unit's base], and the [colonel/captain] ordered the soldiers to [abandon/evacuate] their camp, but they were unable to [avoid/evade] the German troops. Isabel and John were close, and the knowledge that he was under the control of the Nazis put [severe/intense] strain on her and her parents. They reacted by sending as many letters and [parcels/packets] via the Red Cross as they could, so that he would not [die of/perish due to] starvation. John was kept as a prisoner for the duration of the war, and was shipped back to England after Germany's surrender. He arrived home [paler/pallid] and much thinner than he had been, and took quite some time to get better. He never went back to work and suffered from bouts of [melancholy/panic] attacks. John initially lived almost as a hermit, eschewing the world of mortgages and profit margins, and renting a small log cabin near [Brighton/Hastings]. He later took up the [trombone/trumpet] and started to compose [songs/music], eventually achieving moderate success.

SECTION 9

Fortunately Dame **I**sabel was not obliged to forf*ei*t her Oxford scholarsh*i*p – the offer was st**i**ll available after the war, and she commenced her studies that year. Her chosen subject was anthropology. Her choice of subject may seem odd g**i**ven that her major field of expertise came to be natural h**i**story, but through the study of anthropology some of the talents that later made her so well-known *e*merged. Her project on the g**y**psies of the county has been [widely/*i*mmensely] admired; she used a st**i**mulat*i*ng yet factual style, and [observed/not*i*ced] the aspects of the surround*i*ngs that [bypass/*e*lude] most people s**i**mply due to the fact that they [overlook/*i*gnore] them. The project d*i*scussed how, though generally lack*i*ng conscious [spite/mal*i*ce], local people had pushed the g**y**psies to the [edge/marg*i*ns] of v**i**ll*a*ge life.

*I*n**i**tially, while at Oxford, a great deal of Dame **I**sabel's spare time was spent among the members of the Oxford Union, [thrash*i*ng out/d*i*scuss*i*ng] questions of Free W**i**ll as against [Fate/pre-d*e*term*i*nation], the mer*i*ts of moral**i**st*i*c d*i*ctatorsh*i*p, and the prem*i*se that fam*i*ne lies at the root of human confl*i*ct. The early signs were that she might take after her father and end up as an MP, or potentially make use of her journal**i**st*i*c talents.

SECTION 10

However, her third year of coll*e*ge saw a marked change. Her outlook altered and she d*i*sm**i**ssed the cut and thrust of the student d*i*scussions as, to quote her, "not s**i**lly or [banal/stup*i*d], but ideal**i**st*i*c and [fundamentally/*e*ssentially] unhelpful and [self-absorbed/*i*ndulgent]". Rather than talk*i*ng about [fundamental/bas*i*c] ideals, she start*e*d to cons**i**der **i**ssues at a real**i**st*i*c everyday level. She campaigned v**i**gorously for [rais*i*ng the standard of/*i*mprov*i*ng] the somewhat [dirty/squal*i*d] and occasionally [rundown/verm*i*n-*i*nfest*e*d] [student/coll*e*ge] accommodation at St H**i**lda's, while maintain*i*ng [low rental costs/a reasonable tar*i*ff] for those on a tight budg*e*t.

A focus on the d**i**fference she could personally and real**i**st*i*cally make to the people around her was a crucial feature of Dame **I**sabel's approach. Her 1978 book 'Thoughts on Everyday Life' [makes clear/*e*laborates] her ideas. The book offers an unusually personal contemplation of everyday [morals/eth*i*cs], and may yet prove to be her most cons**i**derable legacy. She would not have liked to be termed a woman on a m**i**ssion, but, though not pompous**,** antagon**i**st*i*c or too moral**i**st*i*c, she certainly tried to use what **i**nfluence she had to get across her [point/mess*a*ge], that **i**gnorance and **i**ndolence are not [good/val*i*d] reasons for do*i*ng noth*i*ng.

SECTION 11

A further [central/*i*mportant] aspect of Dame **I**sabel's time at Oxford was her unfulf**i**lled prom*i*se to marry R**i**chard Hawk*i*ns, a former naval capt*ai*n. Their romance had [prospered/flour*i*shed] through their mutual membersh*i*p of the [hockey/tenn*i*s] and [lacrosse/badm*i*nton] clubs, and M**i**ster Hawk*i*ns went on to have a career as a successful d**i**plomat. Although both of them [kept quiet/were d*i*screet] and wouldn't goss*i*p about their break-up, the main facts are nevertheless known to the publ*i*c. As **I**sabel's career took off, her fiancé would not accept that she could sat*i*sfy the duties of an ambassador's wife while travell*i*ng the world on her own, and

asked her to g**i**ve up her chosen profession. Dame **I**sabel d*i*sm**i**ssed the idea, and would not adm**i**t that her own career was [secondary/*i*nferior] to that of her fiancé, lead*i*ng [finally/*e*ventually] to their spl**i**t. Her actions made her a f**i**gurehead for equal rights campaigners**,** due to her oppos**i**tion to the conventional [primacy/prec*e*dence] of men's careers. However, she was wary of opportun**i**st*i*c equal rights campaigners who might draw general conclusions about marr*ia*ge from her own personal ordeal. When asked on Woman's Hour, **I**sabel answered that each [case/marr*ia*ge] should be analysed separately; sometimes [there were/_] obvious solutions or compromis*e*s [that could be made on both sides/would *e*merge], while, for other couples, one partner's w**i**sh*e*s might have to supercede the other's, creat*i*ng an awkward [and sometimes unworkable state of affairs/state of affairs such that the marr*ia*ge cannot be salv*a*ged].

SECTION 12

Shortly after she had left Oxford, Dame **I**sabel's first 'break' came. She was [asked/*i*nvit*e*d] to take the post of anthropolog*i*st on a voy*a*ge to Sumatra. Her role was to study the tribal peoples that they came across, while her colleagues surveyed the creatures and plants of that section of the [jungle/trop*i*cs]. The survey took eight months, and showed **I**sabel the joys of travel. A short documentary about the tr**i**p was comm**i**ssioned by the BBC. The programme, "A Tr**i**p to the Jungle", was a success, and Dame **I**sabel's talent for [convey*i*ng/*i*mpart*i*ng] facts to the general publ*i*c was apparent.

She was then asked to do a series of programmes on the peoples of the Andes. Although again mak*i*ng a success of present*i*ng, **I**sabel realised that anthropology was not where her true passion lay. Her autobiography states, "I found that, when writ*i*ng my diary, more than half my time was spent on [long accounts of/lav*i*sh pass*a*g*e*s] about the wildlife I came across, and realised that I was most [absorbed/*i*nvolved] when writ*i*ng about the marvellous variety of flora and fauna. I thought there was [an audience/a mark*e*t] for programmes that portray the wonderful flora and fauna of the world."

SECTION 13

Dame **I**sabel was clearly right, but the proposal was hard to pursue further. The size of the cameras and their f**i**lms made the kind of on-the-ground wildlife photography that we see today out of the question. Her **i**nput to the jungle documentary had largely cons**i**st*e*d of photographs and p**i**ctures of tribal peoples, and question*i*ng of the scholarly professors who had studied them. Her first wildlife programme took the same approach, but [aimed/*i*ntend*e*d] to **i**nterview people who had direct hands-on knowl*e*dge of their subjects. The focus was on tales from those who had personally tracked or handled the creatures to be d*i*scussed. She was able to draw out *i*nvolved answers from even the most t**i**m*i*d of her guests; the first ep*i*sode contained an excellent sequence about [the crafty strategies of a dolph*i*n, who learned to knock the buck*e*t of f**i**sh from the keeper's hands/a dolph*i*n who *e*volved a crafty tact*i*c whereby he could knock the buck*e*t of f**i**sh from the keeper's hands]. Other highlights were the first ever p**i**ctures of *I*mperial Salamanders and Galapagos peng*ui*ns.

SECTION 14

Dame **I**sabel's next step was what made her such a well-known f**i**gure: she chose to make wildlife programmes for a young audience. The job matched her friendly persona and lively **i**nterest perfectly. The affection that many feel for her and the grief that they must feel on her death no doubt stems largely from their fond childhood memories of watch*i*ng the TV, gr**i**pped by the [actions/ant*i*cs] of some strange creature. Although birds and mammals were the most obvious [subjects/top*i*cs] to appeal to youngsters, Dame **I**sabel was able to talk passionately about plants and **i**nsects by focuss*i*ng on their most amaz*i*ng and unusual properties. The carn**i**vorous Venus Fly-trap was [pleasurably/sat*i*sfy*i*ngly] [gruesome/horr*i*d], while the record-break*i*ng [conker/turn*i*p] was awe-*i*nspir*i*ngly huge, and the [ugly/gobl*i*n-like] Giant Cockroach was just agreeably [gr**i**sly/d*i*sgust*i*ng]. Her programme 'Nature for the Young' drew an [eager/av*i*d] audience and made nature fun for a whole generation of young people.

Her later programmes, for both adults and youngsters, covered such diverse subjects as the h**i**story of woodland, people's use of plants through the centuries, and 'Bugs Around the Home'. The latter made uncomfortable view*i*ng for those who prefer to *i*gnore the carp*e*t-moths, earw*i*gs, woodlice and mites that lurk throughout our homes! The project that gained her the most plaud*i*ts was 'Nature Around Br**i**tain', a programme that gave a brief overview of a range of landscapes across the country, and the wildlife they support.

SECTION 15

When 'Nature Around Br**i**tain' was broadcast, green **i**ssues [were not/had not yet *e*merged as] a major concern for most people, and Dame **I**sabel was one of the first to attempt to [convey to the publ*i*c/*i*nform the publ*i*c of] the [harm/dam*a*ge] that human actions could cause for the future of the [world/plan*e*t]. She raised concerns about pollution, such as [factory waste/waste flu*i*ds] enter*i*ng the network of streams and r**i**vers, [beer bottles/plast*i*c pack*e*ts] and other [waste/rubb*i*sh] dumped under hedgerows, and the **i**mpact of [air pollution/traff*i*c congestion] on plant growth.

As technology progressed, her programmes were able to show more and more spectacular [photography/**i**m*a*g*e*s], and for her most recent series, 'The Nation's Gardens', the cameras were able to show what goes on underneath the soil of people's gardens. They could also fast forward [ivy/clemat*i*s] scrambl*i*ng [up a wall/over the trell*i*s], and speed up the life cycles of garden flowers such as [asters/tul*i*ps], [sunflowers/jasm*i*ne] and delph**i**niums, and produce gruesome close-ups of aph*i*ds feast*i*ng on a rose-bush. These, her final programmes for the screen, were a testament to what made her so special: [the power/us*i*ng the power of lang*ua*ge] to make everyday objects seem **i**nterest*i*ng.

SECTION 16

Dame **I**sabel also wrote copiously about her travels. She and her companions frequently v**i**s*i*t*e*d what was virg*i*n land for Europeans, so the wildlife f**i**lm*i*ng was often accompanied by novel d*i*scoveries. The [malodorous/putr*i*d] ac*i*d propelled from the armp*i*t of the macaque monkey, and the notion of ur*i*ne as a s**i**gnal to warn off rivals were first documen*e*d by her colleagues. Dame **I**sabel also has several journal

papers to her name, although the less formal accounts of her trips are the most widely read. The diaries of her East [Australian/African] tour are the most popular, as much for their stories of the curious antics of her fellow travellers as for their tales of the natural world. The story of the seizure of their [gear/luggage] by armed [rebels/bandits] brandishing muskets as the party filmed at a [peaceful/tranquil] waterhole cannot fail to [recreate/evoke] an atmosphere of danger and adventure, and tells us of the very real risks that sometimes accompanied their trips. Dame Isabel was not merely an [adept/accomplished] broadcaster who had [an eloquent/a fluid] and [clear/lucid] writing style, but was also a woman of great [bravery/courage].

SECTION 17

Though Isabel made many trips to distant countries – from the [dark/humid] jungles of the [equatorial countries/the tropics], and the [scorched/arid] deserts of [Australia/Africa], to the [vast/immense] snow plains of [Siberia/the Antarctic] – she was passionately fond of Britain and upbeat about the future of British wildlife. Through her work for the Conservation Trust, she sought to pass on her zeal to other people. "The pleasure gained from time spent observing nature cannot be overstated," she once wrote. "Whether a [blackbird/robin] outside the window, a pied wagtail at the [local park/sewage works], or an eagle viewed from the [top of a crag/summit of Ben Nevis], the sight creates a sense of wonder." Rather than spending her free time walking round [London/Dublin], [viewing/inspecting] the fashionable shops of [Rome/Paris] or even touring the canals of [Bruges/Venice], Dame Isabel would rather take a break from her [packed/hectic] schedule and unwind by striding the gentle hills of southern England, or [fleeing/vanishing] to the peace of the Welsh hills or the isolation of [Dartmoor/Bodmin] Moor.

Dame Isabel occupied one further public position: her role as Head of the Humanist Association, a not-for-profit organisation that promotes moral values that aren't based on faith or superstition. Brought up as a strict Methodist, her convictions were first shaken at the age of twelve, when she saw the carnage of a car crash at first hand, and started to ponder the random nature of death.

SECTION 18

She [went to/worshipped] at chapel occasionally while at St Hilda's, but, as her workload grew heavier and other commitments grew to be more time-consuming, so her attendance dropped. Although not outspoken about her lack of faith, Dame Isabel was widely assumed to be an atheist, and her high public profile led to her being asked to support the foundation of the Humanist Association. Speaking on Desert Island discs, Dame Isabel said her loss of faith was a conclusion reached through the gradual realisation that the traditions and [dogma/doctrine] of those whom one looks up to are not always appropriate any longer – as science moves forwards, so one's views should change. "I realised," she said "that the rituals of the Christian faith should be questioned as rigorously as are the moonlit druid customs that take place on [June the 21st/the summer solstice]."

Billy [Turnpike/Morris], the current Head of the Association, said that Dame Isabel would be greatly missed. Doctor [Turnpike/Morris] [hopes/intends] to lead the funeral: "We hope to make the funeral [a joyful/an upbeat emotional] occasion rather than a mournful one," he said. "We have put together an assortment of her much-

loved [songs/mus*i*c] to be played. One of the less well-known facts about her was her love of food, and so we have organised a meal that comprise*s* some of the d**i**sh*e*s she brought back from her travels. They range from the fairly standard [pesto/rock*e*t] and ham salad to her special [Bulgarian/*I*talian] [cheese/ol*i*ve] and [tomato/garl*i*c] bread, and the Madagascan [lemon/mandar*i*n] and banana muff*i*n topped by l**i**quor*i*ce allsorts to round off. We th**i**nk she would have approved wholeheart*e*dly," he f**i**n*i*shed.

SECTION 19

Dame **I**sabel was g**i**ven her damehood shortly after her eighty-first birthday. The honour recognised her many years of valuable broadcast*i*ng, part**i**cularly for young people, and also acknowl*e*dged her efforts to br**i**ng an appreciation of the outdoors to a wider population. The national newspapers, down to the most hardened pund*i*t, confirmed that the award had been well-earned, and Dame **I**sabel was very moved by the honour, call*i*ng the award a splend*i*d and memorable occasion.

She cont**i**nued to work throughout her early eighties, but at age eighty-five her doctor pronounced a verd*i*ct of bowel cancer, and she was unable to work after her treatment commenced. She never off**i**cially announced that she had stopped work, and may have hoped to cont**i**nue even longer, but that was not real**i**st*i*c; the treatment, [comprisi*n*g/*i*nvolv*i*ng] surgery and chemotherapy, was unsuccessful and the cancer spread. Nevertheless, she was not too [downcast/morb*i*d], and [was grateful for small mercies/d*i*splayed *i*mmense cour*a*ge], writ*i*ng that "at least I shall st**i**ll be able to appreciate the beauty of the world when the time comes for me to die". She passed away last Friday night at a hosp*i*ce near her childhood home.

SECTION 20

Dame **I**sabel's final broadcast goes out on radio next year. A radio series called 'Access for All' turned out to be her final project, and might be said to demonstrate many of her values. The programmes [outline/d*i*scuss] the problems faced by the least well-off c**i**ty-dwellers when they try to access open country, demonstrat*i*ng the tortuous routes, awkward timetables and [huge/ep*i*c] numbers of t**i**ck*e*ts and perm*i*ts that must be acquired to get away from the rat-race. For those who must overcome lang*ua*ge barriers, and who struggle to afford [clothes/pock*e*t-money] for their k**i**ds, a tr**i**p to the country cannot be their main concern. The programmes cons**i**der the solutions that have been or could be worked out to overcome these problems, and the [strategies/tact*i*cs] that local groups can make use of to get access to the bas*i*c help they need. As such, the programmes address Dame **I**sabel's [amb**i**tion/targ*e*t] to get more people to care about the natural world, as well as demonstrat*i*ng her talent for *i*nspir*i*ng people to th**i**nk that they can real**i**st*i*cally achieve change for themselves.

Her death must, of course, be a loss to the country, but the constant flow of tr**i**butes to her life and work show that her sp**i**r*i*t l**i**ves on through the support and ass**i**stance that she gave to so many. Fortunately, a q**ui**ck [search/rumm*a*ge] through the BBC archives shows that many of her early documentaries survive, and the cont**i**nued success of her books and programmes guarantees her a last*i*ng legacy.

APPENDIX B

EXPERIMENT 1: TRAINING – COMPREHENSION QUESTIONS

The correct answer is presented on the left.

1. What was Dame Isabel most well-known as?
   a. a naturalist b. a geographer
2. What did Dame Isabel try to collect from the tree?
   a. a robin egg b. a pine cone
3. What did Dame Isabel try to save from building work?
   a. a meadow b. a wood
4. What type of animal did Dame Isabel try to adopt?
   a. a dog b. a cat
5. Who owned a farmhouse near the Davis' holiday home?
   a. Beatrix Potter b. E. Nesbit
6. Which college awarded Dame Isabel a scholarship?
   a. St. Hilda's, Oxford b. Girton, Cambridge
7. What skill did Dame Isabel develop while working at the Met Office?
   a. map reading b. typing
8. What is Dame Isabel's brother called?
   a. John b. James
9. What did Dame Isabel study at Oxford?
   a. anthropology b. biology
10. What did Dame Isabel campaign for at Oxford?
    a. better accommodation b. mixed sex colleges
11. On which Radio 4 programme did Dame Isabel comment about marriage?
    a. Woman's Hour b. The Today Programme
12. What was the title of Dame Isabel's documentary about Sumatra?
    a. A Trip to the Jungle b. An Island Adventure
13. Which creatures were filmed for the first time in Dame Isabel's wildlife documentary?
    a. Galapagos penguins b. polar bears
14. What was Dame Isabel's programme for young people called?
    a. Nature for the Young b. Wildlife for Kids
15. What did Dame Isabel bring to public attention?
    a. environmental issues b. rural poverty
16. Which of these was first documented by Dame Isabel's colleagues?
    a. urine use as a signal b. tool use in apes
17. Which organisation was Dame Isabel in charge of?
    a. the Humanist Association b. the Methodist Syndicate
18. Where was the muffin recipe from?
    a. Madagascar b. Ecuador
19. How old was Isabel when she was awarded her damehood?
    a. eighty-one b. sixty-nine
20. What will Dame Isabel's final broadcast discuss?
    a. access to the countryside b. the nation's gardens

APPENDIX C

EXPERIMENT 1: STIMULI FOR LEXICAL DECISION TASK

Words				**Non-words**			
Unstressed keywords	**Stressed keywords**	**Filler words**					
napkin	kindergarten	folder	undertake	plogwer	jumberhortong	prelder	geederstoom
province	convince	collide	lullaby	jockruff	banhunth	kapike	vooragay
varnish	technician	mellow	confession	verpludge	tof'forshen	moopoo	jugpesses
erase	syrup	spiny	computer	opproshen	storren	zainlun	rumwayler
rustic	sticky	tidal	bandwagon	huglan	yackree	whyjup	bungepoffer
anvil	villain	hamper	newspaper	enkop	gronjoo	hompel	nogtapput
bodice	disco	grovel	ratio	gladutch	sasschoo	krovnet	glachiup
comic	Michaelmas	freezer	meteor	groovunge	hockerpass	meezen	yeetium
crevice	viscous	rapture	rational	breeshern	kwaytle	paptule	brushenat
ermine	mineral	mammoth	percussion	arhime	wonager	rammet	semozzup
implore	impulse	heirloom	pavlova	oyflair	omgroon	ayrook	petjeeple
frolic	flicker	gamble	affluent	hertuck	kruffer	scrumbet	efflium
carriage	abridge	fatal	angina	zommudge	petroodge	jaytuff	hestynem
placid	assiduous	duvet	bonanza	drassut	ossamputuss	loovuck	follantem
Justin	tinsel	keyboard	medium	fogyor	pantrun	bulgap	weedeeat
tacit	citrus	cocktail	questionnaire	greejum	kogroopadair	tospun	potcheneer
fetish	optician	cocoa	tremulous	glabbesh	abtershup	poeyop	kwomulut
fabric	bricklayer	plughole	subconscious	gumfroll	pollgryer	gupkyre	obsonten
maverick	ricochet	shoddy	troublesome	hooferum	broddumplay	shraptee	gubbleton
lurid	riddle	blockade	volcano	tryklun	wairdle	plakyle	polcarto
tonic	nicotine	translate	notation	kwootang	yeckerloon	yanprote	snopaction
gremlin	linear	vampire	glamorous	flatchon	hommerack	vokploot	skrematous
assassin	syndrome	ponder	lethargy	effkoona	webtroom	yopken	merthagen
vomit	submit	startle	hydrogen	tromwer	fro'grepp	painster	horgessen
avarice	crystal	spiral	gallery	adajun	hoopstrel	narstal	kalluntee
culprit	pretty	measure	formula	bulgon	screlly	neeplore	jormuna
inert	index	denture	eavesdropper	op'plerk	kankecks	boction	poonbrugger
stolid	Lydia	cowboy	drudgery	holprud	joopia	valposh	vedgator
pontiff	Tiffany	Jacob	cultural	yugtroff	graffapee	kevuck	bolturood
osmosis	sister	Marcus	calculator	aswarkuss	frester	harspog	felbremator
rabid	forbidden	ladle	corroborate	tralluck	demudgen	kaygel	serranberoot
bailiff	proliferate	cartoon	absolutely	pieflup	ralufteroon	yerstub	rabsoenighton
polish	abolition	bonfire	corporation	pogflash	ebutrarsion	hobnird	grooparytel
circuit	kitten	thunder	xylophone	terflan	wotkrun	muffdo	boopapok
blemish	omission	apply	laborious	warklep	apairtrock	umquee	hendarian
adage	prodigious	cattle	librarian	oggum	gosslajous	pattub	jayprariun
terrapin	pinafore	cranky	perforation	fettadell	sedapew	janpee	ferpotriytion
Darwin	windy	parade	propaganda	blorkum	wandoo	komayj	workagondo
engine	ginger	gyrate	segmentation	eprew	jendeem	hayvite	lartemation
refurbish	bishop	bumper	congratulate	har'kruntle	borlom	fomhush	stonpetronate

APPENDIX D

EXPERIMENT 1 : PRACTICE MATERIALS

Practice training story

Jonathan Peters was twenty-one when he applied to work at London Zoo. He was taken on as a Junior Keeper, and looked after the Reptile Centre. The job was generally fairly mundane – the snakes had to be fed several times a day, and there was always mess to clear up and errands to run. However, he liked to talk to the people who came to view the reptiles, and was happy to show them all of the most unusual creatures. The main attractions were unquestionably the dangerous Peruvian python and the well-camouflaged chameleons.

Practice comprehension question

Where did Jonathan work?
a. The Reptile Centre b.The Snake Sanctuary

Practice lexical decision stimuli

Practice words	Practice non-words
panther	groopolls
galaxy	soojunk
radio	wollyferbet
casserole	rattabong
rectangular	vecktrunkulum

APPENDIX E

EXPERIMENTS 2A AND 2B: TRAINING STORY

Key : N= Noun

V= Verb

SECTION 1

Mary(N) stands(V) nervously on the step(N), and tentatively rings(V) the old-fashioned doorbell(N). The door(N) is opened(V) by an elderly lady(N) with long, white hair(N) hanging(V) to her waist(N). She smiles(V) kindly.

"Well, my dear(N)", she says(V). "Shall we start(V) with a cup(N) of tea(N)?"

And so Mary's(N) first-ever interview(N) begins(V). She has just started(V) to write(V) for her school magazine(N) and has come(V) to meet(V) Dame Isabel(N)-Davis(N), the famous naturalist(N).

Colourful books(N) line(V) the walls(N). Mary(N) has never seen(V) so many before. Isabel(N) observes(V) her interest(N). "What you might term(V) 'the natural world(N)' has always fascinated(V) me," Isabel(N) explains(V). "I collected(V) shells(N) as a child(N). My poor mother(N) thought(V) that I would fall(V) to my death(N) while pursuing(V) some rare specimen(N). Not without reason(N) either," she grins(V) mischievously. "I survived(V) several close shaves(N). I collected(V) and classified(V) fossils(N) too, and identified(V) ferns(N). So these books(N) represent(V) a lifelong interest(N)."

Mary(N) feels(V) nervous. Isabel(N) pioneered(V) wildlife television(N), created(V) the first broadcasts(N) aimed(V) at young children(N), and led(V) attempts(N) to alert(V) the British public(N) to the environmental damage(N) that human actions(N) can cause(V). Mary(N) asks(V) who inspired(V) her.

The great lady(N) considers(V). "I don't think(V) I can pick(V) one specific person(N). Rather, a great many people(N) supported(V) and encouraged(V) me. My parents(N) treated(V) me identically to my brother(N). We attended(V) the best schools(N) they could afford(V), and they regarded(V) my academic achievements(N) as just as important as my brother's(N). I didn't feel(V) like a pioneer(N), probably because the people(N) who cared(V) for me held(V) liberal views(N). I just took(V) what seemed(V) the obvious course(N) and applied(V) to university(N). In retrospect(N), I realise(V) how well I was treated(V). "

"I thought(V)," says(V) Mary(N), who wants(V) to prove(V) that she has done(V) her research(N), "that you were prevented(V) from starting(V) university(N)."

"True," Isabel(N) replies(V). "But that difficulty(N) was caused(V) by the war(N). My father(N) wanted(V) me to attend(V) university(N) – I'd been awarded(V) a scholarship(N) at Oxford(N) – but I volunteered(V) for the services(N) instead. My father(N) argued(V) that anyone(N) could volunteer(V) for the services(N), but very few people(N) could study(V) for a degree(N) at a prestigious university(N). He intended(V) well, but I felt(V) strongly that I should serve(V) my country(N). I

helped(V) to manufacture(V) munitions(N) for two years(N), and then worked(V) for the navy(N). Fortunately the university(N) allowed(V) me to start(V) my course(N) after the war(N) had finished(V)."

"Obviously not everyone(N) escaped(V) so lightly," Isabel(N) continues(V). She looks(V) momentarily sad. "The war(N) merely disrupted(V) and delayed(V) my life(N), but my brother(N) experienced(V) much worse. He served(V) just a few months(N) before he was taken(V) prisoner(N). The experience(N) changed(V) him radically. He didn't return(V) to his job(N) and lived(V) the rest(N) of his life(N) with our parents(N)."

Mary(N) sips(V) her tea(N) because she doesn't know(V) how to respond(V). She looks(V) at her notepad(N) for inspiration(N). "What achievement(N) do you recall(V) with most pride(N)?" she asks(V).

Isabel(N) smiles(V) slightly at the sudden change(N) of subject(N). "My work(N) supporting(V) schools(N)," she says(V) without hesitation(N). "Fortunately, I don't need(V) to worry(V) about my finances(N) and so I helped(V) to establish(V) schools(N) in many of the remote rural areas(N) that I visited(V). I find(V) it hugely rewarding to return(V) to observe(V) the difference(N) that the schools(N) have made(V) to the locals(N). I like(V) to see(V) the children(N) laugh(V), smile(V), learn(V) and generally value(V) their education(N). I hope(V) the establishment(N) of these schools(N) will continue(V) to benefit(V) children(N) for years(N) to come(V)."

SECTION 2

"Did you always want(V) to make(V) wildlife films(N)?" Mary(N) asks(V).

"No. Actually, I rarely watched(V) television(N) because my parents(N) didn't own(V) one. My career(N) really just unfolded(V) naturally. As I expect(V) you know(V), I studied(V) anthropology(N) at university(N). My first film(N) investigated(V) tribal people (N) in Sumatra(N). The trip(N) was proposed(V) mainly for academic research(N), but the documentary(N) turned-out(V) really well. Then I presented(V) another series(N) which documented(V) the indigenous people(N) of the Andes(N). I enjoyed(V) the travel(N) and loved(V) the adventure(N) – we became(V) the first Europeans(N) to contact(V) some of the more remote tribes(N). However, I realised(V) that the wildlife(N) fascinated(V) me even more than anthropology(N) so I began(V) to study(V) it seriously.

SECTION 3

"You know(V), I've spent(V) much more time(N) as an academic(N) than I've spent(V) as a broadcaster(N)," says(V) Isabel(N), "but my documentaries(N) made(V) a much bigger impact(N) than my academic papers(N)!"

Mary(N) hadn't realised(V) that Isabel(N) had ever published(V) any papers(N). “What did you write(V) about?” she asks(V) curiously.

“My first papers(N) described(V), illustrated(V) and classified(V) the new species(N) we encountered(V) on our expeditions(N). But my recent research(N) has focussed (V) more on how environmental changes(N) affect(V) and potentially endanger(V) various species(N). I expect(V) I can think(V) of a British example(N)... Well,

recently I investigated(V) whether noisy traffic(N) affects(V) whether birds(N) find(V) a mate(N). Traffic(N) reduces(V) the audible range(N) of their songs(N), you see(V).”

SECTION 4

"I expect(V) you found(V) your expeditions(N) exciting," Mary(N) says(V). “Did anything(N) dangerous happen(V) to you?"

"Well, once a black panther(N) scared(V) us badly while we filmed(V). It sprang(V) from nearby trees(N) and tried(V) to attack(V) us. My scariest moment(N) occurred(V) in Kenya(N), though. The Kenyan authorities(N) gave(V) us permission(N) to film(V) at a waterhole(N) where elephants(N) frequently drank(V). However, the local tribespeople(N) reacted(V) aggressively. Perhaps they hadn't been warned(V) of our arrival(N). They threatened(V) us with guns(N) and stole(V) all our equipment(N), which included(V) all our films(N) and food(N). Fortunately we managed(V) to find(V) transport(N) to the nearest town(N) fairly quickly. However, that experience(N) stands-out(V) as very unusual. Most of the people(N) we met(V) reacted(V) in a friendly way(N) to our presence(N)."

SECTION 5

"Which programmes(N) did you most enjoy(V)?" Mary(N) asks(V).

Isabel(N) responds(V) swiftly. "Definitely my first programmes(N) for children(N)," she says(V). "I exploited(V) children's(N) fascination(N) with funny, gruesome and bizarre creatures(N) to make(V) some really exciting programmes(N). For example(N), I featured(V) orchids(N) that dissolve(V) and consume(V) their prey(N), birds(N) that migrate(V) and manage(V) to navigate(V) huge distances(N), and poisonous insects(N) that sting(V), paralyse(V) or otherwise disable(V) much larger creatures(N). Producers(N) rather neglected(V) children(N) in the early days(N) of television(N), and I think(V) this series(N) helped(V) to change(V) things(N)."

"Didn't you ever want(V) children(N)?" Mary(N) asks(V). Then she realises(V) she has just asked(V) a very personal question(N) – not one of the questions(N) that her teacher(N) approved(V). She blushes(V). "I mean(V), you know(V), if you enjoyed(V) the programmes(N) so much ...".

SECTION 6

Fortunately Isabel(N) smiles(V) at the question(N). "My career(N) always took(V) precedence(N)," she replies(V). "I decided(V) not to marry(V) for the same reason(N). My fiancé(N) refused(V) to believe(V) that my career(N) could exist(V) alongside his job(N), and I refused(V) to leave(V) my profession(N) so we separated(V)."

"Did you make(V) the right decision(N)?" Mary(N) asks(V), feeling(V) like a proper journalist(N).

"Maybe, maybe not." Mary(N) looks(V) unhappy at this answer(N) so Isabel(N) explains(V) her point(N). "Life(N) involves(V) so many decisions(N). I've enjoyed(V) my career(N). I hope(V) I've made(V) a difference(N) to some people(N).

I don't regret(V) my decision(N), but I expect(V) I'd have enjoyed(V) life(N) even if I'd made(V) other choices(N)."

SECTION 7

"You say(V) you want(V) to make(V) a difference(N)," says(V) Mary(N). "Do you really think(V) that your campaigns(N) on environmental issues(N) influence(V) how people(N) behave(V)?"

"Yes, definitely. I often receive(V) letters(N) asking(V) for advice(N) about environmental lifestyles(N). Some people(N) ignore(V) the green message(N), but as we contact(V) more and more people(N), our chances(N) improve(V). Recently I've campaigned(V) to highlight(V) difficulties(N) with access(N) to the countryside(N). This obviously affects(V) principally urban areas(N) and poorer people(N).

To persuade(V) the public(N) to care(V) for the countryside(N), they need(V) to experience(V) it. If people(N) realise(V) that their actions(N) affect(V) places(N) they care(V) about, they may change(V) their habits(N). They might recycle(V) more waste(N), or save(V) energy(N), or buy(V) more local produce(N).

SECTION 8

"I suppose(V) that, for a young person(N), all these concerns(N) seem(V) old news(N). But the extent(N) of our impact(N) on the environment(N) was discovered(V) only relatively recently. Before the eighties(N) one rarely heard(V) such views(N) expressed(V)."

"Do you intend(V) to retire(V)soon?" asks(V) Mary(N) – Isabel(N) has reached(V) her eighties(N) and has continued(V) to work(V) well beyond the usual age(N) for retirement(N).

"I certainly don't intend(V) to retire(V) yet," Isabel(N) says(V) decisively. "However, I've witnessed(V) enough friends(N) lose(V) their health(N) to know(V) that schemes(N) about when to retire(V) sometimes go(V) awry. At the moment(N) I still feel(V) the need(N) to work(V), and particularly to promote(V) my charity(N), which funds(V) research(N) into ecology(N). I suppose(V) perhaps the focus(N) of my work(N) has shifted(V). I conduct(V) less of my own research(N) and instead support(V) the work(N) of other people(N)."

SECTION 9

"Do you think(V) you'll make(V) any more programmes(N)?" Mary(N) asks(V).

"Possibly," answers(V) Isabel(N). "My most recent series(N) will start(V) on television(N) shortly, so I'll wait(V) to see(V) the reception(N) of those programmes(N) before I decide(V) whether to commit(V) to another series(N)."

"What does the new series(N) deal(V) with?" asks(V) Mary(N).

"People's(N) gardens(N)," answers(V) Isabel(N). "That may seem(V) unexciting, but actually it results(V) from a wider project(N) which aims(V) to increase(V) public awareness(N) of local wildlife(N).

SECTION 10

A few years(N) ago, perhaps such a programme(N) would have been seen(V) as dull, but new technology(N) enables(V) us to go(V) to new places(N), and film(V) on a totally different timescale(N). We can follow(V) a mole(N) underground, or film(V) inside a robin's(N) nest(N), or observe(V) trees(N) burst(V) into blossom(N). When you remember(V) the enormous cameras(N) we used(V) on our first expedition(N), the progress(N) seems(V) absolutely incredible."

Mary(N) scribbles(V) in her notebook(N). She thinks(V) she's written(V) enough. She stands(V) up. "I enjoyed(V) the tea(N) enormously. Your answers(N) will make(V) a really interesting story(N). I'll send(V) you the article(N) when I've written(V) it!"

APPENDIX F

EXPERIMENTS 2A and 2B: LABELLING STRESSED VOWEL IN NOUNS AND VERBS

The table below summarises the criteria used to label the stressed vowel in nouns and verbs in the training story and sentence beginnings in Experiments 2a and 2b, prior to lengthening or shortening the section between the two labels.

Preceding segment	**Criteria for labelling start of stressed vowel**
nasal	Onset of vowel formant structure as indicated by onset of complex waveshape.
stop	Onset of periodicity after burst
/r/	End of F3 and F4 transition
/l/	Abrupt increase in amplitude of formants
/j/	Auditory percept
/w/	Increase in amplitude of F3 and F4 and end of F2 transition
/f/ /θ/ /s/ /ʃ/ /tʃ/ /h/(voiceless)	Onset of periodicity
/v/ /z/ /h/(voiced)	End of frication and abrupt increase in amplitude of formants
Following segment	**Criteria for labelling end of stressed vowel**
nasal	End of vowel formant structure as indicated by end of complex waveshape
stop	Onset of stop closure indicated by end of periodicity or abrupt reduction in amplitude
/r/	Start of decrease in F3 and F4 frequency
/l/	Start of rise in F3 frequency or abrupt decrease in amplitude of formants. For coda clusters (e.g., help, felt, held), the end of F3 transition was used in order to avoid the labelled section being too short
/f/ /θ/ /s/ /ʃ/ /θ/ /tʃ/	End of periodicity
/v/ /z/ /ð/	Onset of frication. Or, where frication overlaps with end of periodicity, end of clear formant structure for F2, F3 and F4

Notes:

pursuing, news, views	Labelled section included /j/
pioneered, realise	/ɪə/ was modified
created	/eɪ/ was modified
reacted	/æ/ was modified
fiancé	/ɒ/ was modified
ˈ*television, teleˈvision*	Labelled vowel was dependent on stress in particular token

APPENDIX G

EXPERIMENTS 2A AND 2B: TRAINING – COMPREHENSION QUESTIONS

Sections 1 and 2
What did Isabel do during the war?

Sections 3 and 4
What did Isabel study at university?

Sections 5 and 6
What did Isabel feature in her television programmes for children?

Sections 7 and 8
What was Isabel's recent campaign about?

Sections 9 and 10
What type of research is funded by Isabel's charity?

APPENDIX H

EXPERIMENT 2A: EXPERIMENTAL SENTENCES

Experimental sentence beginnings were recorded with the noun and verb sentence endings in the table below. **Bold** type indicates the keywords, and *italics* indicate which version was truncated to create the sentence beginning used in the experiment.

Sentence beginning	**Noun ending**	**Verb ending**
the Indian **braves**	are dancing round	*a hostile crowd*
the college **grants**	are just needing substance	*a short leave of absence*
the French **cook**	attested it was pepper	*a tasty sort of pepper*
the police **report**	ashames the force	*a shameful crime*
the primrose **flowers**	are just like the bluebell	*ahead of the bluebell*
the official **lies**	are patently nonsense	*about the corruption*
the major **fights**	are described with trouble	*a determined struggle*
the Russian **pines**	*are good species for here*	a good deal for his wife
the community **notices**	are torn at the corners	*a tern in the borders*
the graduate **plays**	*are staged in the park*	a staid little guard
the government **houses**	*are packed with thieves*	a pack of thieves
the factory **fires**	*are burning in the warehouse*	a quarter of the workforce
the animal **calls**	*are declining*	a companion
the army **supplies**	are hardly safe	*a party game*
the leopard **spots**	*appeal to her taste*	a teal in the lake
the guest **reviews**	*are hardly inspired*	a hostel in Spain
the British **show**	ashames the nation	*a shameful station*
the chief **demands**	*are changed with ease*	a change of teas
the dwarf **plants**	*are tiny trees*	a tiny tree
the terrorist **states**	*are changed by policy*	a change in policy

APPENDIX I

EXPERIMENT 2B: EXPERIMENTAL SENTENCES

Experimental sentences were recorded with the verb- and noun-compatible endings below, prior to truncation and merging using STRAIGHT.

Sentence beginning	**Ending**
He'll make Norma/Norm a **bet** if	Jack takes part in the race
They made Suzannah/Suzanne a **drink** whenever	she was ill
We made Elena/Elaine a **pack** because	there were lots of ham sandwiches
They made Diana/Diane a **print** in case	the computer broke down
I'll make Paula/Paul a **deal** if	Tim doesn't appear soon
He made Maria/Maria **sign** because	Dan was deaf
He made Daniella/Danielle a **train** because	*he was so enthusiastic
They'll make Helena/Helen a **budget** so	she can afford her weekly shop
We made Georgia/George a **model** so	there'd be no question of favouritism
She made Joanna/Joanne a **dress** so that	she could appear in public
They made Philippa/Philip a **drum** because	they were short of percussionists
I made Peter/Pete a **plan** although	Paul had attempted it already
I made Christina/Christine a **snack** whenever	Joe's energy levels were low
He'll make Eva/Eve a **bowl** if	Pete says yes
They made Gina/Jean a **file** since	they knew she liked paperwork
I made Louisa/Louise a **spy** even though	she was very reluctant
He made Tricia/Trish a **copy** since	Tim had made so many mistakes
They made Jonah/Joan a **promise** so	Madelaine felt better
They made Carla/Carl a **report** because	Tim's behaviour was so bad

* 'he' used instead of 'she' to avoid coarticulation that would cause problems for truncating at end of 'because'

APPENDIX J

EXPERIMENTS 2A AND 2B: FILLER SENTENCE BEGINNINGS FOR SENTENCE COMPLETION TASK

Experiment 2a	**Experiment 2b** (where different from 2a)
Polly dropped the catch	
A castle stood on the hillside	
In the village	She knows the villagers gossip
Luckily	Luckily Emma remembered that
During the speech	
They began	They began to laugh because
They flew	
She believes that the boy	
He gave us a	He gave Jim a
An enormous	
He vowed to be back	
Apparently the exams	
Maybe	Maybe Amy had told him because
He laughed as he listened to the	
Terry did his homework	
The directory listed six	
She giggles	She giggled because Neil said
Her flatmates	She complained about them even though
He fears that they will find	He fears that Tony will
Throughout the afternoon	
Mark wondered whether	
In the tropics,	
She said	Mel said she'd rather eat snails than
Obviously the man sat on	
He saw the stone	
A modern picture hangs above	
After a bad start to the season	
I climb a mountain	
An informative display explains	
He cuts the sandwich	
He asked to go	
Martha started to write	
In the park	
Nevertheless	She shocked John when
For the last five years	
We spent ages	
Funny people	They wanted to fight but
Presumably the coast erodes slowly	
Towards the end of the concert	
On sunny days	
Everyone agreed to say	

They wanted	I took the blame even though
He insisted on	
A horse grazed in the	
Amanda worried that	
Perhaps	Perhaps John won because
He jumped the gate	
He likes bananas	
I asked why	I asked why Susan had left in the middle of
However	Whenever Vi asks for help
They plan to allow	They intend to allow
After a terrible week	
She wants to drive	
Of course	Henry phoned the police because
Bad language	
Through the binoculars	
Jeremy demonstrated that his theory	
Jean typed the letter	Elsie typed the letter before
Ollie threw a ball	
I hate the creepy	
They grinned	
My car	My car stalls whenever
I think that she	
They hoped	The vicar hoped that Millie would sell lots of
When we painted the bedroom	
James wished	
She forgot her keys	
Jenny realised that	
Marcus kicked the wall	
I managed to smile	
Three years ago	
His lessons	His lessons became a nightmare because
A tiny little spider ran across the floor	
A large crack appears in the ceiling	
She campaigned to release	
She grew two inches	
They promised to leave	They agreed to leave
A drop of wine	
She crossed the floor	She crossed the floor to join
Because of the delays	

APPENDIX K

EXPERIMENT 3: TRAINING STORIES

*Re- p*refixes are indicated by **bold** type, and /ri:/ in stems by *italics*.

STORY 1

Ma*rie* looked up from the magazine she was *rea*ding.

"What's your starsign?" she asked.

"A*ries*," replied Jane. "And I don't think my horoscope is likely to be helpful."

"It's time for a new start," read Ma*rie*. "You must **re**-think your ambitions and **re**-write your *dreams*. The future is bright, but only by **re**-laying the foundations of your life can you hope to achieve what is possible."

"Well, that's ve*ry* specific," said Jane sarcastically. "That's *rea*lly going to help me sort my life out. What does yours say?"

Mar*ie* **re**-located the page. "Leo," she read. "Your naturally care*free* temperament will need to be controlled if you are to make the most of opportunities in the month ahead. Suppress your instinctive response and **re**-weigh your options. A more dis*creet* approach will lead to in*tri*guing results…"

"I could write those things," commented Jane. "All they ever do is **re**-hash the same stuff over and over. I wonder how much you get paid. Maybe I *rea*lly should **re**-think my ambitions and decide to become a fortune teller. Anything to help **re**-pay the student loan – and the rent!"

"I think you'd be good at it, " Ma*rie* smiled, "but I doubt there are many vacancies. But se*ri*ously, you should think about **re**-joining the temping agency you signed up with last summer. Tempora*ry* work's the best route to a decent job nowadays. A way of getting your foot in the door at the best companies."

"I could," admitted Jane unenthusiastically. "It just seems such a waste of my deg*ree*. Last summer I just ended up **re**-stocking the shelves at the station*ery* store. But I guess you're right. I've **re**-read the job section about *three* times, and there's nothing remotely plausible."

"Exactly," a*greed* Ma*rie*. "You need to **re**-prioritise. Stop looking for your ideal job, and just look for a job that's good enough. You need to **re**-direct your efforts."

"You sound like a horoscope," said Jane.

STORY 2

A light *breeze* rustled the leaves, and the *stream* babbled gently down the hillside. "I think we should **re**-locate to Devon," stated Janet firmly.

"And I think we should **re**-trace our steps," said her husband. "It looks like rain. You'd have to get used to rain if we **re**-located to Devon," he added.

They made their way back to the holiday cottage, and arrived just as the first drops of rain began to fall. A row of ever*green trees* stood beside the driveway, and a *cree*per climbed romantically up the front wall. They rushed inside, and Janet tried to **re**-light the fire, which the children had allowed to burn out.

James immediately turned on the laptop. "Work," he said *brie*fly. "My editor wants me to **re**-write the first part of chapter 4."

"I thought you'd already **re**-drafted that about fifty times," complained Janet.

"That's exactly why," explained James. "I just can't get it right. The more I **re**-write it, the worse it gets. But I haven't looked at it for a couple of weeks, so maybe it'll work this time."

"Well, make sure you save it," warned Janet. "Remember last time when the computer crashed and ev*ery*thing had disappeared when you **re**-started it. You had to **re**-type the whole thing."

"As if I could forget." The *screen* glowed brightly, and James was soon engrossed in his story.

Janet sighed a little as she went into the kitchen to **re**-heat the leftovers from yesterday in the shiny new microwave. They had rented the cottage once before, many years ago, and the expe*ri*ence had been quite different. It had been Janet's idea to return, in an attempt to **re**-discover that peace and tranquility they'd found before. But things had been different from her *dreams*. Some things were an improvement – the whole place had been **re**-wired for a start, so they didn't have to wor*ry* about the children sticking their fingers in dangling plug sockets. But she wasn't sure that internet access was *rea*lly a necessity in a holiday cottage. Somehow, they hadn't managed to **re**-*cre*ate that sense of isolation and *free*dom from the stressful *re*alities of their ev*ery*day lives.

STORY 3

Te*re*sa was pretty satisfied with the way it had turned out in the end. Ev*ery*one had thought it was a disaster at first – the fire had destroyed most of the galle*ry*, and they had had to close. But the insurance money had come through quickly and the main galle*ry* had been completely **re**-built within the year. The con*crete* floors had just been **re**-carpeted in a sensible dark-*green*, and the walls were freshly painted in *cream*. The rooms at the back had been less badly damaged, and had simply been **re**-decorated. Te*re*sa's own office had been **re**-painted in a se*rene* blue. She felt fully **re**-charged and ready for anything.

It would be a while before they **re**-constructed their customer base, but she had no doubt that the regulars would eventually return to the galle*ry*, and there were some exciting contempora*ry* exhibitions coming up that should bring in newcomers. Only 2 hours to go now until she could *greet* the first customers. Te*re*sa **re**-polished her desk, which was already glowing mahogany-brown, and **re**-positioned the pot-pour*ri* so that it would instantly give a tasteful impression to visitors but wouldn't obscure the computer *screen*.

The galle*ry* shop would open in a few days time. They'd had to completely **re**-stock, because ev*ery*thing had been so blackened by the smoke that it was impossible to sell. Te*re*sa herself had acquired a few bits and pieces from the de*bris* that was left after

the fire. There were a few **re**-prints of famous Cornish artists which had been in decent condition, and an elegant copy of a figu*rine* from their latest exhibition. She had placed these around her office, and was ex*treme*ly pleased with the result.

Altogether, it felt like the galle*ry* was undergoing a kind of **re**-birth, a chance to start afresh. OK, she wouldn't want to expe*ri*ence devastating fires too *fre*quently, but this one had proved a *real* opportunity.

STORY 4

"Well, how was your day?" asked I*rene's* mother. "Did you enjoy your shopping *spree*?"

I*rene* put her bags down on the kitchen table. "It was OK. I wouldn't exactly call it a *spree* though – I haven't got enough money for a *spree* until I finish **re**-paying you for the damage to the car. But fashion is so *freak*-ish at the moment that I don't *rea*lly care. Eve*ry*one's wearing dunga*rees*, which I think look *cree*py.

"Did you get your shoes **re**-soled like I suggested?" said her Mum.

"Yes, that was a good idea. I didn't know cobblers still existed, " said I*rene*. "He said he could **re**-sole my sandals too if I wanted. It's so nice to be able to keep my favourite shoes instead of chucking them out."

"I've been getting into this secondhand stuff too," said I*rene's* Mum. "I went on *Free*cycle today. It's amazing what you can get. I've requested a big ceramic plant pot so I can **re**-pot the bay *tree*. The fern needs **re**-potting too, but there wasn't anything suitable. I've also offered that old desk of yours for collection. It's a bit of a mess, but someone might want to **re**-condition it, and it's been hanging around uselessly in the garage for ages. Did you see the film that you wanted?"

"No, we were a bit unlucky," said I*rene*. "It was full by the time we got there so we watched Har*ry* Potter instead. Did you know they're **re**-running lots of old movies at the moment? Things like Casablanca and *Brief* Encounter. They've been digitally **re**-processed, and the colours are supposed to be much sharper now."

"So cinemas are getting into secondhand films too!" said I*rene's* Mum.

"They've been into them for years, only not quite so obviously" I*rene* pointed out. "Practically eve*ry* film that comes out nowadays is a **re**-make."

STORY 5

Har*ri*et found moving house a somewhat sur*rea*l expe*ri*ence. It was totally bizarre to think that after she left she'd never **re**-cross the threshold of 2*3 Green Street*, never make another cup of tea on the old gas oven, and never **re**-hang her coat on the Pinocchio-nose coat peg.

Bill had driven off several hours ago, in order to be at the new place when the removal vans arrived. She would **re**-join him later, but being alone for the final few hours was surprisingly emotional, and she couldn't help **re**-living the past as she walked despondently from room to room.

Har*ri*et **re**-checked the house one final time. Had all the boxes been picked up? Was the gas switched off? Had she remembered to **re**-set the burglar alarm and give the code to the new owners?

As soon as she left, it would be the end of an era. It would be hard starting out again, even if Bognor *Re*gis was a nice a*re*a. There would be new neighbours to meet, she'd have to **re**-learn the skill of chatting politely to people she didn't know, and somehow they'd have to **re**-make their lives in the new community.

Still, the new house had potential. Admittedly it was a bit unkempt and run-down, but she was looking forward to making all kinds of improvements. The whole place would need **re**-decorating – the previous owners had a taste for floral wallpaper, and she had some ideas about **re**-doing the kitchen, although that would have to wait until they'd saved up.

But most of all she was looking forward to sorting out the garden. Bill had given her total *free*dom to do what she wanted. She intended to completely **re**-shape it. She'd get rid of the huge fir *tree* that blocked out most of the light, and **re**-plant the area with a few small fruit *trees* and some wild flower seed, to try to **re**-*cre*ate the atmosphere of an old orchard. There was space for a small pond, which she could plant *reeds* around, to give the frogs somewhere to hide. And she planned a typical English coun*try*-house border near the house, with hollyhocks, roses and lupins.

It would be a big change from the small terrace, which they'd had for the past six years. Har*ri*et knew their **re**-location was for the best *rea*lly – the promotion was a fantastic opportunity for Bill – but she couldn't help feeling a bit sor*ry* at what they were leaving behind.

STORY 6

Cooking with kids could be a nightmare, or it could be great fun, Do*reen* mused. Most children viewed it as a *treat*, but this could have ve*ry* different effects on their behaviour. The unruly ones ran all over the kitchen, splashing and **re**-splashing the floor with half-ready cake mixture, making as much mess as possible, while the others were somewhat overawed by the expe*ri*ence, and quietly and carefully obeyed Do*reen's* instructions.

Today's contingent were making *trea*cle tart, strawber*ry* ice-*cream*, and meringues. They'd already had to **re**-start the ice-*cream* after one of the boys spat in it, and then a girl hadn't understood her instruction to *grease* the tart tin, and had *greased* the meringue bowl instead, so she'd had to **re**-wash that. But in general it hadn't been too bad.

The classes she ran for adults were easier in some ways, and harder in others. The adults always wanted to do something sophisticated – she could never save time by **re**-heating yesterday's leftovers as part of a meal. She always used *pre*mium in*gre*dients for the adults too. She'd decided to **re**-stock her cupboards with principally local organic goods after repeated questions about where she sourced her food. People could be a bit too picky sometimes though. She'd once had to **re**-prepare an entire casserole after one of the clients refused to eat it because it contained New Zealand lamb. Not that it had gone to waste – the local homeless shelter occasionally benefitted from her business – they'd **re**-heated it that evening and no-one from *there* asked her to **re**-source her meat locally!

The key to the adults' lessons was teaching them about herbs. Most of them complained that the basil and co*ri*ander plants they bought from the supermarket never survived more than a few weeks, but ve*ry* few of them ever bothered to **re**-pot them, so it was no wonder that the ex*treme*ly overcrowded little pots never lasted ve*ry* long. Do*reen* liked to emphasise that a source of fresh herbs was indispensable to tasty cooking, and so she recommended **re**-trying with a properly planted window box. She also encouraged them to grow their own salads – the cut-and-come-again varieties were ve*ry* convenient – you just let the leaves grow to about 5 or 10 centimetres, then cut them off, and they would **re**-grow within just a few weeks, ready for the fresh leaves to be harvested again.

But the emphasis for the kids' lessons was enjoyment. If they enjoy cooking food, they'll want to eat it, she'd tell and **re**-tell their parents, when they complained about their children's fussiness. For these lessons, Do*reen* weighed most of the in*gre*dients herself, and gave the children the exciting jobs like stirring, pouring and mixing. Of course, she sometimes had to give the in*gre*dients an extra mix round, or **re**-stir the cake mixture to make sure the flour was thoroughly combined, but, for the most part, the kids had these jobs to themselves. If a child did want to weigh, Do*reen* always let them, though she often tried to **re**-weigh anything that was crucial, without the child noticing.

The messiest bit usually turned out to be the washing up, but the young kids tended to see that as part of the fun, especially as the washers always got to eat the uncooked mixture.

STORY 7

"This government has no commitment to the **re**-distribution of wealth," Bob asserted. "They call themselves Labour, but I've seen no sign of them backing the workers' cause."

Jenny sighed unhappily. Eve*ry*time they went out for a nice meal Bob felt the need to **re**-state his political principles. "They call themselves New Labour now," she pointed out. "And you can't say they've done nothing to help the poor. I mean if you just *read* the local paper, you can find out about that scheme to **re**-house the council tenants from that awful tower block round the corner. And they've done a certain amount of **re**-distribution – you can't say that *free* bus travel for pensioners is a bad thing."

"The gap between rich and poor is wider under this government than ever before, and you're talking about *free* bus travel for the elderly," said Bob incredulously. "Do you know how much the average student has to **re**-pay after *three* years at university? And the politicians that voted for tuition fees got their own university education for *free*. Perhaps the Cabinet **re**-shuffle will bring about a shift in policy direction, but I'm not hopeful – it's years since they came up with a *rea*lly innovative policy. Nowadays all they ever do is **re**-publicise the same ideas under a different headline and hope that we don't notice that nothing's changed."

Jenny a*greed*. "Most people don't *read* the paper properly," she said. "They could literally **re**-print the same sto*ry* and by and large no-one would notice. But you should admit that they've done *some* things for education – there are dilapidated classrooms all over the coun*try* being **re**-built."

"I didn't say it was all bad," Bob pointed out. "For example, I think the way they've **re**-positioned themselves on Europe is strategically quite clever. But if they don't concede some ground to those on the lowest wages, then I think we're in for a season of strikes, and I don't want **re**-live the 1970s, even if you do."

STORY 8

Hen*ry* put the phone down. "Sorted," he said. "Let's hope it all works out this time." He'd just been **re**-booking their holiday in *Greece* after the company they'd *pre*viously booked with had gone bust. "Now I suppose I'd better **re**-book the flights too, but I can do that online."

Bronwen looked up from her sewing. "You always get annoyed when you do things online. Last time you had to **re**-dial about twenty times after the signal failed at the critical moment – it'd probably be quicker to phone."

Hen*ry* shook his head. "It'll be OK now we've got broadband," he said. "And I can compare prices myself online – I found some pretty good deals yesterday, and I'd like to **re**-check them in my own time."

Bronwen shrugged. Experience had taught her that Hen*ry* and computers were not a good combination, but there was no point trying to persuade him to **re**-think his decision. She wasn't surprised when, twenty minutes later, Hen*ry* was complaining because the computer asked him to **re**-confirm his password yet again.

It was always better to give Hen*ry* some *breath*ing space when he was irritable, so Bronwen ignored him and **re**-threaded her needle. She had bought a new pair of trousers at discount price because the hem needed **re**-stitching. You could *fre*quently get good bargains that way, with just minor modifications needed. She'd once got a dress half-price just because the sleeve was slightly ripped – she'd removed both sleeves and **re**-*cre*ated the dress in a strapless style. In fact, she was hoping their *Greek* holiday would be the ideal opportunity to show off her *cre*ative efforts. Assuming that Hen*ry* managed to **re**-locate those good deals on flights, they would be off within the week.

"Finally." Hen*ry* **re**-joined Bronwen on the sofa. "Eve*ry*thing's done. This time next week we'll be **re**-discovering what it feels like having nothing to do except enjoy the sunshine."

STORY 9

The new *priest* was doing his best to in*crease* support for the local church, which could no longer be thought of as the centre of the community. The *pre*vious vicar had done his duty, but nothing more, and so Reverend Johnson felt there was a lot of work to do in the parish.

For a start, he had to **re**-build relations with the local school. The headmaster wasn't keen to a*gree* to any more input from the church than was absolutely necessa*ry*, because the *pre*vious vicar had not been good with children and was a particularly dull *prea*cher. In contrast, Reverend Johnson used to be a teacher, and was keen to make the most of his skills. He had **re**-trained as a vicar through a sense of vocation, rather then because he hadn't enjoyed teaching, and he believed it was possible to *preach* and engage children at the same time.

The congregation had been relatively large about ten years ago, and the new vicar was also trying to **re**-motivate the villagers to attend church more *fre*quently – both the regular weekly service and on special occasions. The *pre*vious incumbent had refused to conduct second marriages, which had caused some hard feelings in the parish. Reverend Johnson was sor*ry* about the situation. He could ap*prec*iate how difficult it must be when one's *creed* went against the wishes of the parishioners, but personally he had no principled objections to **re**-marriage, and indeed, he was glad of the extra income that weddings brought to the church. Nowadays, he performed in*crea*sing numbers of **re**-commitment ceremonies too. He found it touching that couples wanted to **re**-confirm their marriage vows.

The church was sorely in need of funds at the moment. As with so many village churches, it desperately needed **re**-roofing – there were several leaks. Ideally, the eastern wall needed **re**-facing, but that was a big job and could probably wait. On a smaller scale, Reverend Johnson was also trying to get the wall **re**-pointed before sections of it decayed even further, and, as a wheelchair-bound parishioner had pointed out, the path to the main door needed **re**-surfacing.

He recognised that it would be a long-term project, and a lot of fundraising would be needed. Luckily the local Scout and Guide groups had already a*greed* to help, and Reverend Johnson intended to en*treat* the WI to **re**-direct their talents towards the Church Roof fund. He felt he'd made a good start, and soon the church would have a big role in **re**-generating the old community spirit.

STORY 10

The local newspaper would be more exciting than usual this week. The big sto*ry* was the trial of Martin Crouch, the alleged bigamist and fraudster. This was a **re**-trial, the first one having been abandoned when members of the ju*ry* had been discovered communicating out of court. The bigamy side of things was great for a personal interest sto*ry*, while the technicalities of his **re**-selling scam, which involved selling the same goods to va*ri*ous different clients, before declaring bankruptcy would interest those who liked to *read* about crime,

That was the front page, but there were some decent articles inside this week's issue too: otters had been **re**-discovered in the local branch of the River *Bre*don after an absence of twenty years, which would please environmentalists. Another positive sto*ry* was the opening of the **re**-styled town community centre, which had been been **re**-named 'The Community Zone' in an attempt to appeal to the younger generation. A bowling alley had been added, and the fittings could now be cunningly **re**-positioned for va*ri*ous activities – small stage performances, film nights, wedding parties, etc.

Parents would be interested in the scandal about school grades. After it was revealed that a teacher had failed to teach the appro*pri*ate text to English A-level students, the entire class's marks had been **re**-graded, so those who had *pre*viously achieved Cs were **re**-classified with As, Ds were **re**-classed as Bs, and so on. There were bound to be a lot of opinions about this, which could only be good for the letters to the editor next week.

The sports section was somewhat more disappointing – the key football matches had been **re**-scheduled due to the appalling weather, the local darts team had lost all their matches, and the school championship hockey match had resulted in a draw, so there

would be a **re**-play in the summer. There was a definite need for something that would **re**-kindle the interest of local sportsmen. Perhaps they could *cre*ate a sto*ry* by **re**-working the old controversy about the rugby field being being sold off by the council to make way for the ma*ri*na **re**-development over in Allen*tree*.

APPENDIX L

EXPERIMENT 3: TRAINING – COMPREHENSION QUESTIONS

Story 1
Where did Jane work last summer?

Story 2
What did Janet do in the kitchen?.

Story 3
What colour was the new carpet?

Story 4
What did Irene's mother offer for collection on Freecycle?

Story 5
Why are Harriet and Bill moving?

Story 6
What does Doreen say is indispensable to tasty cooking?

Story 7
What does Jenny say that New Labour has done for pensioners?

Story 8
Why were Bronwen's trousers at a discount price?

Story 9
What job did the Reverend Johnson have before becoming a vicar?

Story 10
What is the new name for the town community centre?

APPENDIX M

EXPERIMENT 3: FILLERS FOR INTELLIGIBILITY-IN-NOISE TASK

Target words are in **bold**. The righthand column shows the SNR at which the sentences were presented.

Filler sentence	**SNR (dB)**
He said the **shop** closed early that day	-8
The ingrained **prejudice** was hard to fight	-8
Barbara left the **company** after she'd lost faith in the executives	-9
The twins were **overexcited** about the upcoming excursion to Gloucester	-2
Daniel was ordered to **purchase** six clipboards	-6
The **terrain** was somewhat rocky and Mike fell over	-5
They disliked the soldiers who had been sent to **rescue** them	-4
There's no such thing as a **bad** pint of beer	-7
Parking on **residential** roads was banned in Tavistock	-9
He insisted on **listening** to each tape twice	-10
The enthralling **lecture** held their attention	-4
The television programme made him **incandescent** with rage	-3
Tracy owned five **iguanas** and a snake	-10
Those curtains are made from heavy **crimson** velvet	-8
As soon as it was over, Sam **started** to rant uncontrollably	-7
Uncertainty causes disruption among **students**	-5
The gang **intended** to raid the supermarket	-1
Jack was incapable of **holding** down a job	-3
The **purple** solution was far too potent	-5
The turquoise **teddy** bear was vile	-4
The stars shone **brightly** over the beach	-7
The protesters threatened to **riot**	-9
She wouldn't eat because the food tasted **insipid**	-5
Finding a word to rhyme with **helicopter** was difficult	-10
They tried to **dissuade** him from studying quantum physics	-9
Southampton was the **safest** harbour on the south coast	-8
Kate liked to rifle through her sister's **wardrobe**	-2
The bleak **wilderness** thrilled her imagination	-1
Oliver's mother was **dreadfully** shocked by his misbehaviour	-10
The population had been **displaced** due to famine	-6
The **rumours** were totally groundless	-4
Inattentive **pupils** are given lunchtime detentions	-7
The high **exchange** rate put a strain on the Turkish economy	-9
They were right to **ration** butter in the fifties	-1
He thought the apples would be **ripe** soon	-3
It was **predicted** that the rain would be torrential	-8

Manfred was **fascinated** by the swinging pendulum	-6
His **hobby** was yelling at strangers	-1
His **appointment** as managing director puzzled them all	-4
The hikers **desperately** needed to rest	-2
Portuguese sailors dance the **hornpipe** better than French ones	-2
His inability to cook was a severe **disadvantage**	-5
The flight to **Toronto** was cancelled due to ice on the runway	-5
Fortunately Sally had escaped to **Albania** just in time	-10
They liked to wrap **presents** for their little sister	-4
Each **October** Luke was forced to rake up the dead leaves	-6
These **partnerships** never work	-8
The stray dog was too **aggressive** for them to keep	-1
Nick was informed of his **punishment** the next day	-6
Monday's coach **journey** proved surprisingly enjoyable	-1
They wanted to rent the first floor **apartment**	-3
Elsie was always **misplacing** her spectacles	-9
The skeleton had lain there for six **thousand** years	-10
William **distrusted** his wife because of her unbelievable excuses	-3
Ten sponge cakes were **left** on the stall at the end of the sale	-7
His taste in music was **impeccable**	-3
A police **cordon** surrounded the scene of the accident	-6
She totally **mistimed** her lines	-2
Annabel was known to be a **shoplifter**	-2
Their **ambition** was to climb in the Andes	-7

APPENDIX N

EXPERIMENT 3: PRACTICE SENTENCES FOR INTELLIGIBILITY-IN-NOISE TASK

Practice sentence	**SNR (dB)**
Cushions make a room seem cosy on cold nights	-2
The aphids were especially bad this summer	-4
Chilli and tomato are a great combination	-6
It helps if you're willing to dance	-6
Melinda hid her fear of spiders to avoid looking like a wimp	-6