UNIVERSITY OF LONDON THESIS

Degree PhD Year 2006 Name of Author EVANS, B. E.

COPYRIGHT

COPYRIGHT DECLARATION

LOANS

REPRODUCTION

Plasticity in Speech Perception and Production: A study of accent change in young adults

by

Bronwen Gillian Evans

Submitted for the Ph.D. degree

University College London

Abstract

This thesis presents the results of two studies that investigated how listeners adapt to different regional accents within the same language.

Study 1 investigated whether listeners adjusted their vowel categorization decisions when listening to speech produced in different accents of British English. The results demonstrated that some listeners chose different vowels according to the accent of the carrier sentence. The patterns of adjustment were affected by individual differences in language background (i.e., the degree of experience that an individual has had living in multidialectal environments, and whether the individual grew up in the north or south of England), and corresponded to changes in production that speakers make as a result of sociolinguistic factors when living in a multidialectal environment.

Study 2 investigated plasticity in speech production and perception among university students, as individuals change their accent from regional to "educated" norms. Subjects were tested before beginning university, 3 months later and on completion of their first year of study. At each stage they were recorded reading a set of test words and a short passage. They also completed two perceptual tasks; they found best exemplar locations for vowels embedded in carrier sentences and identified words in noise. The results demonstrated that subjects changed their spoken accent after attending university. The changes were linked to sociolinguistic factors; subjects who were highly motivated to fit in with their university community changed their accent more. There was some evidence for a link between production and perception; between-subject differences in production and perception were correlated. However, this relationship was weaker for within-subject changes in accent over time.

List of Tables

List of Figures

Acknowledgments

The research described in this thesis was carried out at the Department of Phonetics and Linguistics, University College London, and was funded by an EPSRC Doctoral Training Award.

This thesis would not be what it is without the help, support and advice of many people. I will not mention all of them here, but I would like to take this opportunity to thank those who have been the most influential in helping me achieve what I have over the past three years.

First and foremost I would like to thank my supervisor, Paul Iverson, who has been a great source of inspiration and support over the course of this project: I have thoroughly enjoyed the past three years and learned so much, all of which is testament to the amount of effort and enthusiasm he has also put into the project.

Many thanks are also due to the all those in 'WoHo' for their help, both as subjects in pilot experiments, actual experiments and as sources of advice, support and encouragement. In particular, I would like to thank Stuart Rosen for help and advice with the word recognition experiment, and Valerie Hazan and Jill House, my PhD committee, for very helpful advice on the development of Study 2.

I would like to thank the people who lent their voices to the experiments described in this thesis; Richard Dewire, Graham Hobbs, Graham Fisher and my sister, Natasha Evans. I am also grateful to my Mum, Jackie Evans, and my sister, Alexandra Evans for their help with subject recruitment, and all those who took part in the experiments, particularly the then 6th form students, now university students, of Ashby Grammar School.

Lastly, I would like to thank my family – Mum, Dad, Natasha, Rich, Alex, and, of course, Graham – for their unfailing love, support and encouragement.

For my Grandad

1. Introduction

1.1. Foreword

In multidialectal environments (e.g., large cities such as London) native speakers of different accents interact with one another. Speakers in these environments often avoid variants that are markedly regional or unusual to facilitate communication (Trudgill, 1986) and to appear cosmopolitan (Foulkes and Docherty, 1999). However, they also retain some regional variants to show their allegiance to particular social or geographical groups (e.g., Foulkes and Docherty, 1999; Trudgill, 1986; Watt, 1998). For example, speakers of northern English who live in southern England typically modify some aspects of their accent in order to fit in with southerners; they change their production of the vowel in words such as *luck* from a high back vowel, [ʊ], to a centralized vowel, [ə], so that it is closer to how southerners produce this vowel (Trudgill, 1986; Wells, 1982b). Yet they also maintain some aspects of their regional identity; they retain [a] when producing words like *bath* just like in their native northern English accent, rather than producing it with the southern vowel [ɑː] (Trudgill, 1986; Wells, 1982b). In order to understand speech, listeners must somehow tolerate or adjust to this phonetic variation.

The aim of this thesis is to investigate how listeners in this multidialectal environment adjust to the phonetic differences between accents. There has been little direct research on this topic, and so the thesis takes as its theoretical motivation two areas of speech perception research that have modeled how listeners might adapt to variability in the speech signal; episodic memory research and models of second language acquisition.

Episodic memory research (see Section 3.2.2; e.g., Goldinger, 1996, 1998; Johnson, 1996) has suggested that listeners should be able to fully adjust to differences between different regional accents as long as they have had experience with that accent. That is, individuals who have had experience with a particular accent may be able to recognize similarly accented speech by mapping it onto stored exemplars produced by talkers of a similar accent. Models of second language acquisition [e.g., Perceptual Assimilation Model, Best (1994, Best et al., 1988, 2001); the Speech Learning Model, Flege (1991, 1995, 2002)] however, have suggested that listeners cannot easily adapt their categorization decisions when listening to non-native speech (see Section 3.2.3). They hypothesize that the acquisition of non-native speech sounds is affected by the listener's native language background, such that he/she will assimilate non-native phonemes into his/her closest matching native categories. Similarly, when adjusting to a non-native accent, a listener may assimilate the non-native phonemes into his/her closest matching native categories without making specific adjustments for the accent of the speaker.

The reported studies focus on the phonetic contrasts between two regional accents, one spoken in southern England, Standard Southern British English (SSBE), and the other in northern England, Sheffield English. These accents differ primarily in their vowel inventories and the lexical distribution of a particular vowel (see Chapter 2). Briefly, speakers of SSBE use the vowel [ʌ] in words such as *luck*, but northern English speakers do not have this vowel in their phoneme inventory; they say *buck* with a higher vowel, [ʊ], such that it becomes a homophone or near-homophone of *book*. Southerners and northerners both use the vowels [a] and [ɑː] but with somewhat different lexical distributions; words such as *bath*, *dance*, and *ask* are produced with [ɑː] by southerners, but with [a] by northerners, even though most words that have these vowels do not differ between accents (e.g., *bad* using [a], and *bard* using [ɑ:]).

In summary then, the aim of the experiments presented in this thesis was twofold. First, the experiments aimed to determine if listeners adjusted their best exemplar locations to match talkers of a non-native accent. Second, the experiments examined whether the ability to adjust to a non-native accent was affected by individual differences in their accent background (i.e., if they had been born and raised in the north or south of England), or the type of experience listeners had had with an accent (i.e., if they had had experience of living in a multidialectal environment).

1.2. Overview

Chapter 2 gives a detailed description of the types of possible differences between accents, and describes the differences that were investigated in this study. The chapter also gives a detailed description of the two accents that formed the basis of the experiments conducted in this thesis – SSBE and Sheffield English.

Chapter 3 reviews previous work that has examined how listeners adapt to variability in the speech signal. No previous work has directly addressed how listeners might adjust to phonetic differences between accents within the same language, and so this review focuses on studies of talker normalization, episodic memory research and models of L2 speech perception that have investigated how listeners adjust to other types of variability in the speech signal. In so doing, the review aims to obtain preliminary ideas and make predictions about how listeners might adapt their speech processing strategies when adjusting to speakers of different accents within their own language and how this might be affected by experience with their native accent. Section 3.3 describes how the design of Study 1 enables these predictions to be investigated and Section 3.4 provides a summary of the hypotheses based on the literature reviewed in Section 3.2.

Chapter 4 presents the results of Study 1. The study comprises two experiments that are presented in Sections 4.3 and 4.4. Section 4.3 describes the design and results of Experiment 1 and briefly discusses the implications of these results. Section 4.4 describes the design and results of Experiment 2. Section 4.5 gives a summary of the results of Experiment 1 and Experiment 2 and discusses the implications of these results for existing theories of speech perception.

Chapter 5 introduces and provides the motivation for Study 2. Section 5.2 reviews related literature and discusses its implications for these experiments. Section 5.3 describes the design of the study and how it addresses the questions raised by Study 1 and the literature review. Section 5.4 summarizes the hypotheses made from the literature reviewed.

Chapter 6 presents the results of Study 2. The study comprises three experiments that are presented in Sections 6.2, 6.3, and 6.4. Section 6.2 describes and presents the results of Experiment 1, Section 6.3 the design and results of Experiment 2, and Section 6.4 the design and results of Experiment 3. Section 6.5 gives a summary of the results and discusses the implications of these results for existing theories of speech perception.

Chapter 7 draws together the results of both Study 1 and Study 2. Section 7.1 summarizes the results, and Section 7.2 reviews the implications of the findings for existing models of speech perception. Section 7.3 presents various explanations for the findings. The chapter concludes in Section 7.4 with a summary of the main findings of the thesis and a brief discussion of their implications for current and future work in speech perception.

2. Standard Southern British English and Sheffield English: Description and Comparison

This chapter focuses on the types of possible differences between accents that a non-native listener might have to adapt to, and describes the types of differences that were investigated in this study. In this regard, the chapter gives a detailed description of the two accents that formed the basis of these studies, SSBE (Standard Southern British English) and Sheffield English.

2.1. Differences between accents

The possible differences between accents can be divided into four main categories; systemic differences, realizational differences, lexical differences and contextual distribution differences.

Systemic differences refer to the differences between the phoneme inventories in different accents, i.e., the presence of a phoneme in one accent but not another. For example, speakers of southern English accents use the vowel [ʌ] in words such as *luck*, but northern English speakers do not have this vowel; they say *luck* with a higher vowel, [ʊ], such that it becomes a homophone or near-homophone of *look*. Thus, northern and southern English accents differ in their phoneme inventory: southern English speakers have the phoneme [ʌ] in their phoneme inventory but northern English speakers do not.

Speakers of different accents may have the same phoneme in their native accent inventory, but these may be produced, or realized, differently. For example, both northern and RP (Received Pronunciation) accents have the phoneme /a/ but in northern accents this is realized as a fully open vowel between front and central [a ɐ] rather than [æ] as in RP (Wells, 1982b).

Speakers may also differ in the phoneme that they use to produce the same word. That is, there may be a difference of lexical distribution for a given phoneme. Probably the most well-known example of this in British English is the difference in the production of words like *bath* and *grass* between speakers of northern and southern English. Southerners and northerners both have the vowels [a] and [ɑ:] in their native phoneme inventories, but they do not always use them in the same words, i.e., they have a different lexical distribution. Words such as *bath*, *dance*, and *ask* are produced with [ɑ:] by southerners, but with [a] by northerners, even though most words that have these vowels do not differ between accents (e.g., *bad* using [a], and *bard* using [ɑ:]).

Accents can also have differences of contextual distribution. That is, both accents can have a particular phoneme in their inventory, but the phonetic contexts in which the phoneme occurs may differ. For example, both Scottish and southern English accents have the phoneme /ɹ/ but in southern English accents /ɹ/ only occurs before a vowel. Southern English accents are thus described as non-rhotic. In contrast, in Scottish English, a rhotic accent, /ɹ/ may also occur after vowels within the syllable. Thus, *cart* is produced as /kɑ:ɹt/ in Scottish English, but /kɑ:t/ in southern English accents.

This study focuses on how listeners adapt to systemic differences (i.e., differences in phoneme inventories) and lexical differences (e.g., /baθ/ vs. /bɑ:θ/ in northern and southern English accents) when listening to a non-native regional accent. These differences were chosen as they form the two most important characteristics that set apart northern English from southern English accents. Specifically, these are;

1. the lack of a phonemic opposition between words such as *book-buck*, *look-luck* in northern English accents (systemic difference – northern English does not have the phoneme /ʌ/)
2. differences in lexical distribution of the phonemes /ɑ/ and /a/. Both northern and southern English speakers have these vowels in their native phoneme inventories, but words such as *bath*, *dance*, and *ask* are produced with [ɑː] by southerners, and with [a] by northerners.

The next section gives a detailed description of the principal vowels and consonants of the two accents that were investigated as representative of these accent groups; SSBE and Sheffield English. The reasons for choosing these two accents were methodological. The main part of the research was carried out in London where one of the dominant accents of the community, particularly in the university environment where the majority of the testing was to be carried out, is SSBE. Sheffield English was chosen for two reasons. Firstly, in order to be able to compare the best exemplar locations chosen by speakers in both a northern and southern English accent in the experimental paradigm used here – the Goodness Optimization Task (Iverson and Evans, 2003) – a speaker who could produce native-like versions of both accents was needed. The speaker recruited who was able to do this was originally from Sheffield. Secondly, the accent of Sheffield is very similar to the accent of the community where the researcher grew up, Ashby de la Zouch, which is in the East Midlands (Wells, 1982b; Hughes et al., 2005). This meant that it was possible to recruit a large number of native speakers of this accent for testing in both Study 1 and Study 2.

2.2. Description of accents

I have used a standardized format of presentation to enable comparison of the principal monophthongs, diphthongs and consonants of SSBE and Sheffield English accents. The format used follows that developed by Wells (1982a) and also used by Docherty and Foulkes (1999).

Transcriptions of the vowel pronunciation norms are presented using the keywords used by Wells (1982b) to characterize the lexical sets of English. The keywords are "intended to be unmistakable no matter what accent one says them in" (Wells, 1982a: xviii), and have become a standard tool within descriptive dialectology. The set of keywords used here is as follows:

KIT	DRESS	TRAP	LOT	STRUT	FOOT
GOOSE	NURSE	FLEECE	FACE	PALM	GOAT
NORTH	PRICE	CHOICE	MOUTH	NEAR	SQUARE
BATH	happY				

The characteristic pronunciations of these keywords are summarized in table form at the beginning of the description. I have then included comments on the keywords in each accent that display interesting features.

The consonant inventories of each accent are described using using small capitals, as for the vowels. Comments are made under the headings STOPS, FRICATIVES, AFFRICATES, NASALS, L, R, and semi-vowels.

2.2.1. Standard Southern British English

The description presented here is based on data presented in Wells (1982b), Tollfree (1999) and Hughes et al., (2005). In using the term Standard Southern British English, I am referring to the variety of southern English used predominantly by middle-class speakers in the south of England. This variety can be considered to be similar to what Wells (1982b) refers to as "Near-RP". The accent includes very little in the way of regionalisms and is perceived by the majority of speakers to give an impression of being 'educated', 'well-spoken', and 'middle-class'. It shares some similarities with RP, particularly the variety of RP used by younger speakers. Indeed, distinguishing RP from SSBE speakers is often considered to be a highly contentious and subjective issue (Wells, 1982b).

2.2.1.1. *Vowels*

KIT	ɪ	DRESS	ɛ	TRAP	æ	LOT	ɒ	STRUT	ʌ	FOOT	ʊ
GOOSE	ʉː	NURSE	ɜː	FLEECE	iː	FACE	eɪ	PALM	ɑː	GOAT	əʊ
NORTH	ɔː	PRICE	aɪ	CHOICE	ɔɪ	MOUTH	aʊ	NEAR	ɪə	SQUARE	ɛə
BATH	ɑː	happY	iː,ɪ								

Table I Summary of the principal monophthongs and diphthongs of SSBE and their phonetic realizations

KIT

There is a tendency for traditional /ɪ/ to be replaced by /ə/ in some unstressed syllables. In general, younger people are more likely to have /ə/ and high status speakers are more likely to have /ɪ/ (Hughes et al., 2005). This occurs in the first vowel of the endings *–ility (e.g., possibility)*, *-itive* (e.g. *positive*), *-ess* (e.g. *hopeless*), *-ily* (e.g. *happily*), *-ate* (e.g. *fortunate*), *-ible* (e.g. *visible*), *-em* (e.g. *problem*), *-ace* (e.g. *furnace*), *-age* (e.g. *manage*), *-et* (e.g. *bracelet*).

GOOSE

This vowel is now fronted and is rarely fully back or rounded as is traditional, though the more back variant may still be heard in older speakers. This vowel may also be diphthongal (Hughes et al., 2005).

SQUARE

This diphthong may be monophthongized to [ɛː] (Hughes et al., 2005).

happY

In younger speakers in particular, there is a tendency for this phoneme to be realized as [iː] in final position in words like *city* and *very*, rather than /ɪ/, e.g. *city*, /sɪtiː/ rather than the traditional /sɪtɪ/ (Hughes et al., 2005).

2.2.1.2. Consonants

STOPS, FRICATIVES and AFFRICATES

There are six stops, nine fricatives and two affricates in SSBE (Table 2).

	labial	labio-dental	dental	alveolar	palato-alveolar	Velar	glottal
Plosive	/p/,/b/		/t/,/d/			/k/,/g/	
fricative		/f/,/v/	/θ/,/ð/	/s/,/z/	/ʃ/,/ʒ/		/h/
affricate					/tʃ/,/dʒ/		

Table II Summary of the stops, fricatives and affricates of SSBE. The symbol on the right-hand side represents the voiced phoneme.

The glottal stop [ʔ] can also be considered to form part of the stop inventory in SSBE, though it does not have phonemic status and is largely unnoticed in the accent (Hughes et al., 2005). It is often used to reinforce /p,t,k/ and the affricate /tʃ/

in syllable-final environments, where it precedes the consonant, e.g., picks [pɪʔks]. It may also be used to mark a syllable boundary when the following syllable begins with a vowel, and can stand in place of an intrusive /ɹ/ (see below).

The glottal stop is also used as a realization of word- or morpheme-final /p t k/ when followed by a consonant, e.g., Scotland, wait there. The realization of /p/ and /k/ as [ʔ] is usually restricted to cases when the following consonant has the same place of articulation as that being realized as [ʔ], e.g. back garden.

In younger speakers, there has been an extension of the use of the glottal stop as a realization of /t/ in SSBE. Younger speakers, upper- and middle-class, can be heard using a glottal stop in word final position, before a pause and before a vowel (Hughes et al., 2005). It has been suggested that this is the result of the influence of popular London speech, or 'Estuary English' (Wells, 1984; Tollfree, 1999; Fabricius, 2002).

It is common to find that initial /h/ is dropped when it occurs in unstressed pronouns (e.g., *he*, *her*) and auxiliaries (e.g., *has*, *have*). This is known as H-dropping.

NASALS

There are 3 nasal phonemes in SSBE; bilabial /m/, alveolar /n/ and velar /N/.

Both /m/ and /n/ are normally realized as the labiodental /μ/ where they precede /f/ or /v/, e.g. comfort [kʌɱfət]. Before /θ/ and /ð/, /n/ may be dental and before /ɹ/ it may be post-alveolar.

L

This is the only lateral phoneme in SSBE, though it has 3 allophones; clear [l] which is voiced and found before vowels (and /j/ in older and more conservative speakers in their pronunciation of words like *cure* and *suit*), dark [ɫ] which is also voiced and found after a vowel, before a consonant, and syllabically e.g. in bottle, and voiceless [l] e.g., plate.

Some SSBE speakers use a vowel in place of dark [ɫ] in some environments, e.g. table or beautiful. This process is known as /l/-vocalization. The quality of the vowel that substitutes for /l/ can vary, but most are back and fairly close (Hughes et al., 2005). It is thought that the increase in /l/-vocalization may be the result of the influence of popular London-speech (see Przedlacka, 2002).

R

SSBE is a non-rhotic accent, and so the phoneme /ɹ/ occurs only before a vowel. In SSBE this phoneme has a number of allophones. The most common is the voiced post-alveolar approximant [ɹ]. Following /d/ it is a fricative, [ɹ□], and following stressed /p, t, k/ it is typically devoiced, [ɹ□]. In intervocalic position when the first vowel is stressed, or following a dental fricative, /ɹ/ may be realized as an alveolar tap, [ɾ]. However, this realization is rare in younger speakers (Wells, 1982b). Linking and intrusive R are almost categorical following a non-high vowel (i.e., following /ə ɔː ɑː/)

semi-vowels

There are two semi-vowel phonemes in English; /w/ and /j/. There is a strong tendency in SSBE, as in other English accents, for /j/ to coalesce with preceding alveolar plosives to form an affricate, e.g., *tune*, *dune*. This is known as yod-coalescence and also occurs when a word ending in /t/ or /d/ precedes *you* or *your*, e.g. *would you* [wʊdʒʊ].

2.2.2. Sheffield English

The description presented here is based on data collected by Stoddart et al. (1999) in various parts of the city in 1997 and Wells (1982b).

2.2.2.1. Vowels

KIT	ɪ	DRESS	ɛ	TRAP	æ	LOT	ɒ	STRUT	ʌ	FOOT	ʊ
GOOSE	ʊʉː	NURSE	əː	FLEECE	ɪiː	FACE	eː eːɪ	PALM	aː	GOAT	ɔː əʊ
NORTH	ɔː	PRICE	ɑɪ	CHOICE	ɔɪ	MOUTH	aː, aʊ	NEAR	iə	SQUARE	ɛə
BATH	a	happY	ɪ								

Table III Summary of the principal monophthongs and diphthongs of Sheffield English and their phonetic realizations.

STRUT

[ʊ] is the most common variant and local norm, though [ə] is found in weak position in rapid speech. In educated northern speakers in all parts of the region, this vowel can also be realized as [ə] in both weak and stressed position or as a vocoid somewhat opener than [ʊ] (see Section 2.3 below for a fuller discussion).

FLEECE

[ɪiː] is common amongst all speakers, though this variant competes with [iː] amongst younger speakers.

FACE

[eː] is used by all speakers, sometimes with a slight [ɪ]-glide, [eːɪ]. A half-open and shortened variant [ɛ] is used by some speakers in words like *make* and *take*, but [ɛɪ] is used in words such as *eight, straight, weight.*

GOAT

Both [ɔː] and [əʊ] are possible, though [əʊ] is particularly common amongst female speakers.

GOOSE

As in SSBE, this vowel is no longer fully back or rounded. Instead, this vowel is now fronted and is typically diphthongal.

MOUTH

[aː] is common for all speakers, though [aʊ] is often found in words such as *house*, and [əʊ] in words such as *about, round*, especially amongst older speakers.

happY

This is typically realized as [ɪ], though [e]/[ɛ] are also possible, especially amongst older speakers. There is some evidence that younger speakers are starting to adopt the southern variant [iː].

2.2.2.2. *Consonants*

STOPS, FRICATIVES and AFFRICATES

	labial	labio-dental	dental	alveolar	palato-alveolar	velar	glottal
plosive	/p/ /b/		/t/ /d/			/k/ /g/	
fricative		/f/ /v/	/θ/ /ð/	/s/ /z/	/ʃ/ /ʒ/		/h/
affricate					/tʃ/ /dʒ/		

Table IV Summary of the stops, fricatives and affricates of Sheffield English. The symbol on the right represents the voiced phoneme.

A glottal stop, [ʔ], is used for non-initial /t/, particularly by male speakers; /p, k, tʃ/ are particularly liable to glottal replacement or reinforcement in final position (e.g., *back*, *get*) or at the boundary of two elements in a compound (e.g., *nightlife*). The definite article *the* is often glottalized or preglottalized when following by a dental, and can also be assimilated to [t] e.g., *t'train.*

Medial /t/ is often realized as [ɹ], e.g., *getting* [gE®In].

/d/ is occasionally assimilated either together with or following [n]; e.g., *friends* [frɛnz], *grandma* [gɹanmɑː].

Final stops /t, d/ and fricatives /f, θ, ð/ are often omitted, mainly in function words, e.g., *just*, [dʒʊs], *and* [an, ən], *myself* [maˈsɛl]. This is more common in male speakers.

The dental fricatives [θ, ð] can be substituted with [f, v], e.g., *mouth* [maːf], *brother* [bɹɑvə]. This is particularly common in younger male speakers.

Final /z/ is often devoiced, or lengthened and devoiced. This is more common in older age groups but can be found in younger speakers.

H-dropping (i.e., omission of initial /h/) is fairly regular amongst older speakers, and common amongst younger speakers, though young females normally use initial [h].

NASALS

As in SSBE there are three nasal phonemes in Sheffield English, bilabial /m/, alveolar /n/ and velar /ŋ/.

As in SSBE, both /m/ and /n/ are normally realized as the labiodental /µ/ where they precede /f/ or /v/, e.g. comfort [kʌɱfət]. Before /θ/ and /ð/, /n/ may be dental and before /ɹ/ it may be post-alveolar.

Final *–ing* is commonly realized as [ɪn]; [ɪŋ] is possible, but not frequently used.

L

Laterals are normally clear, but dark [ɫ] is occasionally found in final position, particularly in female speakers e.g., *school* [skʊəɫ], and before final [t,d], e.g., *cold* [koʊɫd].

R

Like SSBE, Sheffield English is non-rhotic, and so the approximant [ɹ] is common for all groups. A tapped [ɾ] is sometimes used by male speakers. Linking and intrusive R are also common for all speakers.

Semi-vowels

As in SSBE there are two semi-vowel phonemes; /w/ and /j/. As in SSBE, yod-coalescence is common following alveolar plosives and when a word ending in /t/ or /d/ precedes *you* or *your*, e.g. *would you.*

2.3. Differences between SSBE and Sheffield English

This section gives a brief outline of the principal differences between the phonemic inventories of SSBE and Sheffield English. Like the descriptions given above, it is not intended to be an exhaustive account of the regional and social differences in pronunciation between the two accents. For a more detailed description and comparison of both accents see Hughes et al., (2005), Foulkes and Docherty (1999) and Wells (1982b).

The /ʌ / vowel

One of the best-known differences between English accents is one of phoneme inventory or systemic difference. Northern accents of English, of which Sheffield English is one, do not have the vowel /ʌ/ in their phonemic inventory. Words that are produced using this vowel in SSBE and other southern English accents, are produced using the vowel [ʊ] in Sheffield English. The vowel /ʌ/ is relatively recent in the history of English, having developed by a phonemic split from the older vowel /ʊ/. Northern English accents, as well as Irish accents, did not take part in this development with the result that in these accents, pairs of words such as *cud-could*, *put-putt* that are distinguished in southern English accents, are not distinguished in the north and midlands of England, and pairs like *blood* and *book,* and *mud* and *look* are perfect rhymes. In northern English accents then, the five-term system of short vowels (/ɪ, ɛ, æ, ʊ, ɒ/) is preserved whilst in the south of England, Wales and Scotland there is now a six-term system (/ɪ, ɛ, æ, ʊ, ʌ, ɒ/).

As the prestige norm (SSBE) exhibits the six-term system, this means that in the north of England there is sociolinguistic variation between the local, less prestigious five-term system, (i.e., with no /ʊ-ʌ/ opposition), and the national, more prestigious six-term system of SSBE (i.e., with an /ʊ-ʌ/ opposition). Consequently, speakers of northern English accents may use a pronunciation style intermediate between their local, northern English accent and SSBE, for example, when interacting with southern English speakers.

There are two phenomena characteristic of this intermediate style (Wells, 1982b). The first is the use of qualities for the STRUT vowel that are distinct from the [ʊ] of FOOT, but that are perceptually distinct from the SSBE /ʌ/. These qualities include a mid back /ʊ/, the unrounded equivalent [ɤ], and more commonly, a mid or half-close [ə] (Hughes et al., 2005). This last-mentioned possibility, a stressed [ə] in STRUT, seems to be particularly common in the variety used by 'educated' northerners (Wells, 1982b), giving pronunciations such as *cup* [kəp] and *brother* [bɹəðə]. As a result of the use of stressed [ə] in STRUT, there is no distinction between the strong and weak form of *but, does, must, us* (Wells, 1982b). Alternatively, educated northern speakers may have stress-sensitive allophonic variation between [ʌ] and [ə], which leads to the appearance of weak forms for words such as *up* and *one*, which have no weak forms in SSBE.

The second phenomena characteristic of this intermediate style is hypercorrection of the vowel [ʊ] in FOOT words (Wells, 1982b). That is, northern speakers may use the vowel they substitute for their [ʊ] vowel in the STRUT words for the FOOT words as well, leading to pronunciations such as *butcher* [ˈbʌtʃə] and *sugar* [ˈʃəgə], for example. It is possible that this is the result of a northern speaker adopting the new, opener, realization for all instances of his/her underlying /ʊ/ category when acquiring a more educated northern English accent (Wells, 1982b: 353).

In summary, Table 5 below shows the typical possible realizations of the STRUT and FOOT vowels in SSBE and northern English accents.

	STRUT	**FOOT**
SSBE	ʌ	ʊ
Northern England [1]	ʊ	ʊ
Modified Northern I	ə	ʊ
Modified Northern II	ə	ə
Hypercorrect Northern	ʌ	ɤ

[1]e.g., Sheffield English

Table V Possible realizations of the STRUT and FOOT vowels in SSBE and northern English accents.

/a/ and /ɑː/

The other well-known feature that distinguishes Sheffield English from SSBE is the use of the vowels /a/ and /ɑː/. Both accents have these vowels in their phonemic inventories but the phonemes have a different lexical distribution. Words such as *grass*, *path* and *laugh* (defined as the BATH lexical set by Wells, 1982b: 133) are produced using a longer, back vowel in SSBE, /ɑː/, but using the short vowel /a/ in Sheffield English and other northern English accents. This difference is the result of the fact that the original short vowel /a/ was lengthened in the south but not the north of England before the voiceless fricatives /f, s, θ/ (e.g., *glass* [glɑːs]) and certain consonant clusters containing an initial /n/ or /m/, though this latter change is less complete, e.g., *dance* [dɑːns] but *band* [bænd] (see Hughes et al., 2005).

/ɪ/ and /iː/

Another feature that differentiates SSBE and Sheffield English is the use of /iː/ and /ɪ/ word-finally, in words such as *city* and *money*. In Sheffield English these are generally produced using the short vowel /ɪ/, whereas in SSBE they are typically produced using the longer vowel, /iː/, particularly by younger speakers. However, there is some evidence that younger speakers of northern English are beginning to adopt the SSBE variant.

[ʔ]

The glottal stop is used in both SSBE and Sheffield English, though its use is typically more widespread in Sheffield English. SSBE speakers may use a glottal stop word-initially before vowels or before certain consonants or consonant clusters (Fabricius, 2002). Speakers of northern English accents such as Sheffield English, particularly young urban working-class speakers, use the glottal stop in more contexts, in particular, as an allophone of word-medial and word-final /t/. However, the glottal stop is appearing more frequently in the speech of younger SSBE speakers (Foulkes and Docherty, 1999, Hughes et al., 2005).

/ŋ/

Sheffield English speakers, along with speakers of other non-standard varieties, generally use /n/ in place of /ŋ/ in the suffix *–ing* e.g., *singing* [sɪŋɪn].

3. Study 1: Vowel normalization for accent

3.1. Overview

In this chapter I review existing literature before outlining the aims and hypotheses of the study and its design.

The literature review in Section 3.2 focuses on how the ability to adjust one's native vowel categorizations to a match a talker of a non-native accent within the same language might be explained. However, little work has focussed directly on how listeners might adjust their phonetic categorizations to match talkers of different accents within their native language. Consequently, this review focuses on related literature that has addressed how listeners adapt to other sources of variability (e.g., when adjusting to anatomical differences between speakers, or learning a second language). Section 3.2.1 outlines why adjusting to differences in a non-native accent differs from adjusting to variability arising from anatomical and physiological differences between speakers. Sections 3.2.2 and 3.2.3 consider how existing theories of speech perception might explain how listeners adjust to a non-native accent.

Section 3.3 focuses on how the ability to learn a foreign language is affected by early experience with a native language, and discusses the effects of age and experience on the ability to learn new phonetic categories. This evidence is then discussed with regard to how age and experience might affect the ability to accommodate to a non-native accent.

Section 3.4 sets out the aims of this particular study and presents various hypotheses based on the literature reviewed. Section 3.5 briefly describes the experimental design and outlines why this task allows the aims of the study to be addressed.

3.2. Adjusting to phonetic variation: how might listeners adjust to phonetic differences between accents?

3.2.1. A comparison with models of talker normalization

Previous research in vowel perception has primarily examined how listeners adjust to the acoustic consequences of anatomical variation, such as the length of the talker's vocal tract and characteristics of their glottis (e.g. Hillenbrand et al., 1995; Ladefoged and Broadbent, 1957; Nearey, 1989). This research has focussed on the development of normalization procedures that aim to describe the perceptual and cognitive adjustments that allow listeners to accommodate to differences between speakers. Listeners have been thought to normalize for anatomical variation by relying on gross acoustic – perhaps language universal – characteristics of the speech signal such as fundamental frequency (e.g. Miller, 1953; Fujisaki and Kawashima, 1968), and the range of formant frequencies used by a speaker (Ladefoged and Broadbent, 1957).

The way in which listeners adjust to differences between accents however, must be different. Studies of vowel normalization such as those mentioned above, rely on positing mathematically predictable relationships between different acoustic factors to explain how listeners adjust to anatomical and physiological differences between talkers. Indeed, vowel normalization procedures such as those developed by Lobanov (1971) and Gerstman (1968) to classify vowel tokens in automatic speech recognition are based on the mathematical relationship between F1 and F2 for a given talker. For example, Gerstman (1968) proposed a procedure that rescales the frequencies of F1 and F2 for each of the vowel tokens for a particular talker. The lowest and highest formant frequencies for each speaker are measured across vowels. These values are then set to 0 and 999 respectively. All other vowels are scaled linearly between these two extremes. It is not necessary to use tokens from all categories of a speaker in order to derive the maximum and minimum value: Gerstman (1968) proposed that /i/ would provide minimum F1 and maximum F2, /ɑ/ can provide maximum F1, and /u/ can provide minimum F2.

Differences in the realizations of phonemes in a different accent cannot be predicted from such basic acoustic factors, though. For example, a northern English speaker when listening to a southern English speaker is not able to predict from the relationship between the talker's F1 and F2 or F0 range that there is a difference in the lexical distribution of the vowels /ɑ/ and /a/. That is, it is not possible for a northern English speaker to predict from such acoustic factors that a southern English speaker uses the vowel /ɑː/ in words like *bath* where a northern English speaker would use a shorter vowel /a/. Equally, a northern English speaker cannot use these acoustic factors to predict differences in phonemic inventory between his/her native accent and SSBE; the relationship between a southern speaker's F1 and F2 range does not allow a northern listener to predict that a southerner uses the vowel /ʌ/ in words like *bud* and *cud* where he/she would use the vowel /ʊ/.

It is thus plausible that the ability to accommodate to differences between accents requires more language-, accent- and phoneme-specific processes. That is, the perceptual and cognitive adjustments that enable a listener to accommodate to, or normalize[1] for differences between accents, must be the result of experience with, and therefore knowledge of, a particular accent. For example, in order to

---

[1] A broad definition of normalization is used here (i.e., the perceptual and cognitive adjustments that allow a listener to accommodate to differences between speakers) rather than a narrow definition that is sometimes used (i.e., a hypothetical perceptual process in which speaker-specific information is discarded; Pisoni, 1997).

accommodate to SSBE *bath*, northerners have to 'learn' that there is a difference in the way in which the phonemes /a/ and /ɑː/ are used in SSBE as opposed to in their native accent, i.e., they have to learn that there is a difference in the lexical distribution of /a/ and /ɑ/ in SSBE and northern English accents. Thus, northerners must 'learn' that SSBE speakers produce words like *bath* and *grass* using the longer, back vowel /ɑ/ that they use in words like *card* and *bard*, rather than the short vowel /a/ that they would use in their native accent.

As described in Chapter 2, there are various types of differences between accents; there are differences in lexical distribution, phonetic realization and phoneme inventory, for example. It is possible that accommodation to these types of differences between accents operates differently, and that certain differences are easier to accommodate to than others. For example, adapting to differences in phonemic inventory may require the acquisition of a new phoneme, as well as learning the words in which this new phoneme is used. For example, northerners do not have the phoneme /ʌ/ in their phoneme inventory. Thus, when interacting with SSBE speakers, northerners must learn that SSBE speakers have the phoneme /ʌ/, as well as learning the set of words in which this phoneme is used. Both northerners and southerners have the vowels /a/ and /ɑ/ in their phonemic inventory but use them with a different lexical distribution. When interacting with an SSBE speaker, northerners must learn that southerners use /ɑ/ in words like *bath* where they would use /a/. One could imagine then, that northerners might be able to adjust to differences in the lexical distribution of /a/ and /ɑ/ in words like *bath* more easily than they might be able to adapt to differences in phoneme inventory where they need to learn that there are differences in lexical distribution and also acquire a new phoneme.

In the following two sections, I examine how existing models of speech perception might explain how listeners adapt to these differences between accents. The first section focuses on exemplar models of speech perception and how these models might produce talker normalization effects that could explain how listeners accommodate to differences between accents. The second section considers how accommodation to a non-native accent might be similar to second language acquisition, and if models of second language acquisition might be able to explain how listeners accommodate to differences in accents.

3.2.2. Exemplar models of speech perception

Exemplar models of speech perception (e.g., Goldinger, 1996, 1998) have theorized that listeners store phonetically detailed memory traces every time they listen to speech. Episodic memory research (e.g., Nygaard and Pisoni, 1998; Palmeri et al., 1993) has shown that listeners encode and retain fine-grained phonetic details of a talker's voice in long-term memory, forming highly detailed speaker-specific representations that are used in subsequent speech perception. For example, Nygaard et al. (1995) investigated the effects of stimulus variability on memory representations for spoken words. A serial recall task was used to study the effects of changes in speaking rate, talker variability and overall amplitude on the initial encoding, rehearsal and recall of lists of spoken words. At short interstimulus intervals (ISIs), talker variability affected memory representations such that there was poorer serial recall for the words from multiple-talker lists than those from single talker lists. However, in the long ISI condition, listeners showed an advantage for multiple-talker lists. It was hypothesized that listeners had associated the distinctive information provided by each of the different voices with the words in the list, and that this had allowed listeners to remember both the word and its temporal position in the list (see also Goldinger et al., 1991). Further evidence to support this view comes from studies that have shown that memory for talker characteristics affects the perception of novel words (Nygaard and Pisoni, 1998), word segmentation (Smith, 2003) and both short-term voice recognition memory and long-term perceptual identification for words (Goldinger, 1996).

Johnson (1997) has suggested that these exemplar representations can produce talker normalization effects, if listeners compare the words that they hear to stored exemplars of speech produced by similar talkers. The ability to accommodate to differences between accents could be viewed as an extreme example of this kind of talker normalization (see Nygaard and Pisoni, 1998; Pierrehumbert, 2000), in that the incoming speech could be compared to stored exemplars of speech produced by talkers of different accents.

For example, individuals who have experience with different British accents may be able to recognize northern-accented speech by mapping it onto similar stored exemplars produced by northern talkers, and recognize southern-accented speech by mapping it onto similar stored exemplars produced by southern talkers. In this respect, there may be no difference in the ability of non-native listeners to accommodate to different types of differences between accents. Thus, provided that they have had previous experience with similarly accented speech, subjects in this study may be able to fully adjust to vowel differences between accents both in terms of differences in lexical distribution and phoneme inventory.

3.2.3. Evidence from studies of second language acquisition

Evidence from cross-language research however, suggests that individuals cannot easily adjust their phonemic categorizations to match the talker, at least when listening to foreign or foreign-accented speech. Two major theories have been proposed to account for the problems that second language learners have in acquiring new phonemic categories: the Perceptual Assimilation Model (PAM; Best, 1994; Best et al., 1988, 2001) and the Speech Learning Model (SLM; Flege, 1991, 1992, 2003). In this section I present a summary of each of these models and consider their relevance and predictions for the multidialectal situation.

In PAM, Best (1994) invokes the concept of cross-language phonetic similarity to predict the relative difficulties that listeners might have in differentiating between non-native phonetic contrasts. Three patterns of perceptual assimilation of L2 segments to L1 phonological categories are described, which are determined by the perceived phonetic similarity of the L1 and L2 segments: (1) as a categorized exemplar of some native phoneme, for which its goodness of fit may range from excellent to poor; (2) as an uncategorized consonant or vowel that falls somewhere inbetween native phonemes (i.e., roughly similar to two or more phonemes); or (3) as a nonassimilable nonspeech sound that bears no detectable similarity to any native phonemes.

Discrimination of a non-native contrast is predicted to depend on how each of the contrasting phones is assimilated. Several pairwise assimilation types are also possible. The non-native phones may be phonetically similar to two different phonemes and assimilate separately to them. This is termed Two Category assimilation (TC). Instead, both may assimilate equally well or poorly to a single native phoneme. This is termed Single Category assimilation (SC). Equally, both non-native phones might assimilate to a single native phoneme, but one may fit better than the other. This is termed a Category Goodness difference (CG). Alternatively, one non-native phone may be uncategorized, as above, while the other is categorized, forming an Uncategorized-Categorized pair (UC). Or, both non-native phones might be Uncategorized (UU). Finally, the two phones' articulatory properties may be perceived to be so different from any native phonemes that they are perceived as Non-Assimilable (NA) non-speech sounds.

The model predicts that discrimination of non-native contrasts will be affected by the relationship of the non-native phones to the listener's native phonology (Best 1994). When the two non-native phones are separated by native phonological boundaries, discrimination will be helped, but when both phones assimilate to the same native phoneme, discrimination will be hindered. However, discrimination for phones that are perceived as nonspeech sounds will be unaffected by native phonology. Consequently, NA contrasts are predicted to show good to excellent discrimination. TC and UC contrasts are predicted to be well-discriminated because the contrasting phones in both cases fall on opposite sides of a native phonological boundary. For CG and SG assimilation types, both phones assimilate to the same native phoneme. If one phone is a good fit and the other is poor, discrimination is predicted to be very good (CG difference), but not as good as in TC contrasts because discrimination is hindered by assimilation to a single native phoneme. In SC cases, both non-native phones are equivalent in phonetic goodness. This is thought to lead to poor discrimination because there is a lack of phonological contrast and a difference in fit with the native phoneme.

Tests of PAM have focussed primarily on the discrimination of non-native consonant contrasts rather than vowels. For example, Best et al. (1988) investigated native English speakers' perception of non-native consonant contrasts in two indigenous African languages, Zulu and Tigrinya. The experiment was designed to test the hypotheses that discrimination should be near-ceiling if the consonant contrast was perceived as phonologically equivalent to a native contrast (TC assimilation type), lower but still good if it was a phonetic distinction between good versus poor exemplar of a single native consonant (CG assimilation type), and much lower if both non-native segments were phonetically equivalent in goodness of fit to a single native consonant (SC assimilation type). The findings supported these hypotheses. For non-native contrasts where listeners perceived a correspondence to a native phonological distinction, discrimination was excellent. When non-native consonants were heard as differing in goodness of fit to a single native consonant, discrimination was still very good but was significantly lower than in the previous condition. Finally, when listeners perceived a non-native contrast as equally good variants of a single native consonant, discrimination was much poorer. One test of PAM using vowels yielded similar patterns of results. Best et al. (1996) tested American English listeners in their discrimination of non-native vowel contrasts in Norwegian, French and Thai. Listeners also performed a keyword identification test for each of the vowel contrasts. The assimilation patterns for the vowel contrasts, inferred from the keyword results, were found to be strongly related to discrimination performance; TC contrasts were better discriminated than CG contrasts, which in turn were better discriminated than SC contrasts.

The Speech Learning Model (SLM) also predicts that the perception of a non-native phonetic category is affected by the L1. The SLM (Flege, 1988, 1992, 1995, 2002, 2003) proposes that the likelihood of new category formation increases as a function of the perceived distance from the closest L1 speech sound. Thus, the greater the perceived distance of an L2 vowel from the closest L1 vowel, the greater the likelihood that a new category will be established for the L2 vowel (Flege, 1995). For example, a native Spanish speaker learning English should be more likely to establish a phonetic category for English /æ/, which differs from Spanish /a/, than for English /i/, which differs only slightly from Spanish /i/.

This hypothesis is supported by evidence from Bohn and Flege (1997). This study examined the perception of English vowels by native speakers of German. Of particular interest was the vowel /æ/ which has no German equivalent. Perception of this new vowel for native Germans was tested in an identification experiment in which a synthetic *bet-bat* continuum was presented to two groups of native German listeners differing in their experience with English; one group were inexperienced and the other were experienced with English. The stimuli varied in duration and vowel spectrum; the formant frequencies for F1-F3 varied from values appropriate for English /æ/ to English /ɛ/ in eleven linearly equal steps, and each of these vowels was presented at durations of 150, 200 and 250 ms. The results demonstrated that the groups differed in the relative effect of the spectral and duration manipulations on vowel identification. Experienced listeners relied more on spectral cues, whereas inexperienced listeners relied more on duration. A control experiment with native English listeners demonstrated that for native listeners spectral information was the most important cue for distinguishing these vowels, but that they relied on this cues more so than the experienced German listeners. This result suggests that experienced German listeners identified the new English vowel /æ/ in a similar way to the native English listeners, suggesting that extended contact with English may precipitate an English-like perception of the /ɛ/ vs. /æ/ contrast. This is consistent with the prediction made by the SLM that for L2 sounds for which there is no obvious counterpart, an L2 learner will eventually establish phonetic categories for a new L2 sound.

If a non-native phoneme is too close to an L1 category though, then the SLM hypothesizes that it will be assimilated into that native category. Like PAM, the SLM hypothesizes that even though a non-native speech sound might be equated with a native sound, a listener may still perceive it to be different from the L1 speech sound with which it has been equated. In this case, it is hypothesized that a "merged" category that reflects properties of both the L1 and L2 phonetic input will develop and replace the original L1 category (Flege, 2002; MacKay et al., 2001).

In summary, both PAM and the SLM predict that non-native phonemes are assimilated to native categories unless they are considered to be uncategorizable (i.e., are categorized as nonspeech sounds) or are perceived to be sufficiently different from a native phonetic category. It is plausible that the perception of a non-native accent operates similarly to these cross-language cases. Although experience with a foreign accent has been shown to improve word recognition performance (Clarke, 2002, 2004), it is possible that, unlike exemplar models have suggested (see Section 2.2.2 above), vowel categorization processes may remain difficult to modify. Consequently, listeners may not be able to easily adjust their categorization processes to match native talkers of that accent. Instead, they may assimilate the incoming speech signal to the phonetic categories of their own native accent categories without making specific adjustments for the accent of the speaker. For example, one could imagine that northern listeners who do not have the southern vowel /ʌ/ in their native vowel inventory might assimilate this vowel to the closest matching native category, e.g., /ə/, when listening to southern or southern-accented speech. If /ʌ/ is judged to be a good example of this category then listeners may assimilate the non-native sound into their native category for /ə/ without making any changes to the category. However, if SSBE /ʌ/ is perceived to be different from the native sound with which it has been equated, i.e., Sheffield English /ə/, then a "merged" category that reflects properties of both the equated native and non-native sounds might develop, as hypothesized in the SLM (Flege, 2002). Thus northern listeners may develop a composite /ʌ/-/ə/ category that replaces their original /ə/ category and which they use to process instances of these two sounds.

3.3. Plasticity in speech perception: Is the ability to adjust to a non-native accent affected by age or experience?

Cross-language research has emphasized age as a factor in the ability to adjust to a non-native language. Studies of second language (L2) acquisition (see Flege, 2003, 2004 for a review) have shown that "earlier is better" as far as learning an L2 is concerned, and that once a critical age has been reached it is no longer possible to learn a foreign language and achieve native-like performance. That is, late bilinguals (individuals who began learning their L2 in late adolescence or early adulthood) usually resemble native speakers of their L2 less than early bilinguals (individuals who began learning their L2 in childhood). For example, early bilinguals generally have milder foreign accents than late bilinguals (Flege et al., 1995), and have been shown to produce and perceive L2 vowels more like L2 native speakers than late bilinguals; early bilinguals produce and perceive L2 consonants more accurately than late bilinguals (Flege, 1992), and recognize L2 words presented in noise more accurately (Meador et al., 2000).

Such age-related effects in second language learning led to the proposal that L2 language learning was constrained by a critical period arising from a loss of neural plasticity – the Critical Period Hypothesis (CPH). According to CPH, the ability to acquire a second language declines with age, as certain mechanisms that are needed for successful language acquisition work less effectively, or become inaccessible beyond a certain age (e.g. Johnson and Newport, 1989; Lenneberg, 1967). Indeed, Patkowski (1990) claims that individuals who begin learning an L2 after the critical period differ in a fundamental way from those who began learning an L2 before the end of this period.

Although CPH researchers have not proposed an exact age at which language learning declines, most agree that the ability to acquire a non-native language is severely impaired by or during adolescence. Patkowski (1990) concluded that the large difference he observed in the foreign accents of speakers who had first arrived in the United States before the age of 15 versus those who had arrived afterwards was due to the passing of a critical period. Scovel (1988) hypothesized that the critical period ends at age 12 years as a result of a decrease in brain plasticity, and DeKeyser (2000) suggested that it ended at age 16-17 years as a consequence of normal neurological maturation.

The ability to adapt to a non-native accent within the same language may be similarly affected. Chambers (1992) studied how six children, aged 9-17 years, who had recently moved from Canada to southern England, adapted their speech production. The results showed that although all children acquired some southern English features, younger children acquired complex, native-like production rules, but children aged 14 and above at the time of emigration did not. Based on this evidence, Chambers concluded that the ability to learn a new accent in one's native language remained intact into early adolescence, but that beyond this period language-learning abilities were impaired such that it was no longer possible to fully adapt to a non-native accent within the same language.

However, there is evidence to suggest that the decline in the ability to acquire a non-native language may not result from a loss of neural plasticity due to the effects of a biologically delimited period. Individuals who began learning their L2 before the end of this critical period have been found to differ from native speakers: Most early bilinguals speak their L2 with a mild but detectable foreign accent (Flege et al., 1997; Piske et al., 2001), and early bilinguals differ from native speakers in the accuracy with which L2 vowels and consonants are produced and perceived (Flege et al., 1999; Piske et al., 2001; Sebastian-Galles and Soto-Faraco, 1999).

Furthermore, evidence from studies of second language acquisition is inconsistent with the existence of a critical period for language acquisition. Flege et al. (1995) investigated English language learning in native Italian speakers who had been living in Canada for an average of 32 years. Subjects recorded five short sentences that were presented to native English speakers and rated for perceived foreign accent. Based on the CPH, one would have expected to see a sharp decline in accent ratings corresponding with the age at which language learning ability became impaired, e.g., during or after adolescence (DeKeyser, 2000; Scovel, 1988; Patkowski, 1990). However, there was no discontinuity in the accent ratings: The ratings decreased systematically as participants' age of arrival (AOA) increased, resulting in a near-linear relationship between AOA and accent ratings.

If the decline in language learning abilities cannot be attributed to a biologically delimited critical period, then how else might they be explained? One explanation is that early experience with a native language constrains subsequent language learning, such that one's native language interferes with the acquisition of non-native speech sounds. Studies of first language acquisition have demonstrated that infants are born with the capability to discriminate all speech sounds, even ones that are not meaningful in their native language (see Kuhl, 2000 for a review). However, by the age of six months, infants show changes in their discriminative abilities. When tested with a phonetic prototype of a native phonetic category, as opposed to a non-prototype from the same category, infants were better able to generalize to other category members (Iverson and Kuhl, 1995, 1996, 2000). This research finding is known as the perceptual magnet effect, because the prototype appears to function as a "magnet" for all other stimuli in the category. The effect is dependent on experience with a specific language: American infants demonstrated the perceptual magnet effect for native phonetic prototypes, but treated non-native prototypes like non-prototypes (Kuhl et al., 1992). These changes in discriminative abilities are thought to result in a language-specific mapping that distorts perception, creating a complex network, or filter through which language is subsequently perceived (Iverson et al., 2003; Kuhl, 2000).

Such early experience is thought to interfere with the ability to acquire a second language, causing second language acquisition to decline with increasing age (see Iverson et al., 2003, Kuhl, 2000). Changes in perceptual processing as a result of early language experience are thought to be self-reinforcing because experience with a native language alters how subsequent speech sounds are perceived. Thus, even though adults may be exposed to the same acoustic distribution of speech sounds as an infant learning the same language, they will perceive them differently due to prior perceptual changes. It is hypothesized that this loss of perceptual sensitivity for non-native phonetic contrasts may be difficult to reverse in adulthood, because perceptual resolution for the types of acoustic variation important for learning a non-native language is reduced. Kuhl (2000) describes this process as one of increasing neural commitment to a particular network structure for analyzing language, and hypothesizes that this has a greater influence with increasing age as a result of self-reinforcing perceptual interference rather than through any biological limitation.

Such an account may help to explain why even though adult listeners can be trained to perceive and produce foreign speech sounds like native speakers, they often do so with difficulty. For example, much research has shown that Japanese listeners have difficulty in discriminating English /r/ and /l/, a contrast that does not exist in their native language (e.g., Goto, 1971; Strange and Dittmann, 1984; Iverson and Kuhl, 1995). Iverson et al. (2003) investigated native versus non-native perceptual differences in the perception of English /r/ and /l/. American English adults and native Japanese adults living in Tokyo rated the acoustic similarity of /ra/ - /la/ stimuli that differed in terms of the F2 and F3 formant frequencies. Multidimensional scaling analyses demonstrated that the perception of acoustic-phonetic information was affected by listeners' native phonetic systems. That is, experience with an L1 had "warped" perception of acoustic-phonetic cues in the L2. Specifically, native English adults showed an increased sensitivity to F3 at the /r/-/l/ boundary, and a reduced sensitivity to F3 differences within the category where the stimuli were identified as either /r/ or /l/. In contrast, Japanese listeners were more sensitive to variation in the F2 dimension, an irrelevant acoustic cue for categorizing /r/ and /l/.

The results suggest that whilst Japanese listeners had developed 'perceptual maps' which were well-suited for Japanese, these maps impeded their acquisition of the English /r/-/l/ contrast. It is possible that the ability to adapt to a non-native accent within the same language may be similarly affected. Early perceptual experience with a particular regional accent may warp perception, such that listeners develop perceptual maps that are tuned to their native accent. These maps may interfere with subsequent perception, such that listeners perceive non-native phonetic contrasts in terms of their native accent, impeding their ability to acquire or adapt to a non-native accent.

3.4. Aims and hypotheses

The aim of Study 1 is to investigate whether listeners adjust their vowel categorization decisions when listening to speech produced in a non-native accent within the same language. Although there has been little direct research on this topic, episodic memory research (see Section 3.2.2; e.g., Goldinger, 1996, 1998; Johnson, 1996) suggests that listeners should be able to fully adjust to differences between different regional accents as long as they have had experience with that accent. Thus, one would expect listeners who have had experience with a non-native accent to choose vowels that match what speakers of that accent would produce when listening to talkers of that accent. In contrast, cross-language research has suggested that listeners cannot easily adapt their categorization decisions to match the talker (see Section 3.2.3; e.g., Flege, 1992, 1995, 2003; Best, 1994; Best et al., 1988, 2001). Models of second language acquisition such as PAM (Best, 1994) and the SLM (Flege, 1992, 1995, 2003) have hypothesized that listeners assimilate non-native phonemes into native categories unless they are considered to be uncategorizable (i.e., are categorized as nonspeech sounds) or are perceived to be sufficiently different from a native phonetic category for a new category to develop. Likewise, it is plausible that listeners in this study may be unable to fully adjust their vowel categorization decisions when listening to a non-native accent. Instead, they may assimilate non-native phonemes into their closest matching native accent categories without making specific adjustments for the accent of the speaker.

3.5. Experimental design

A vowel categorization task, the Goodness Optimization Task (Iverson and Evans, 2003), was used to investigate if listeners changed their best exemplar locations when listening to speech produced in different accents. Listeners found best exemplars for vowels in two different accents of British English, SSBE and Sheffield English. These accents were selected because they exhibit distinct differences in their vowel inventory (Chapter 2). Listeners from different linguistic backgrounds (i.e., northerners and southerners) found best exemplar locations for each vowel in both SSBE and Sheffield English. Subjects thus found best exemplars in both a native and a non-native accent.

The Goodness Optimization Task is similar to the Method of Adjustment Task developed by Johnson et al. (1993), but differs in one important way. In the Method of Adjustment Task (Johnson et al., 1993), listeners chose stimuli from a grid. Listeners were able to hear as many different versions of the stimuli as they wished before deciding which they thought was a best exemplar of the vowel printed on a computer screen. Of particular interest in this study was whether listeners would choose different best exemplars for the vowel in words like *bath*, which is produced differently in northern and southern English accents (see Chapter 2). This difference is highly salient and well-known to speakers of both accents. Consequently, I felt that it was important that listeners were not able to listen to both versions, i.e., they were not able to listen to both [baθ] (Sheffield English) and [bɑːθ] (SSBE) before selecting a best exemplar. As a result, a task was developed in which listeners 'synthesized' a best exemplar for a vowel that was printed on a screen – the Goodness Optimization Task (Iverson and Evans, 2003). Listeners heard a vowel, they gave goodness ratings, and a computer program iteratively adjusted the acoustic qualities of the vowel until a best exemplar was found. The computer program could find a best exemplar in a minimum of 30 trials.

The target word was embedded in a carrier sentence. This carrier sentence was specifically designed to give an overt cue to accent, as listeners were not told on beginning the experiment that they would be listening to either a northern or southern accent. Consequently a longer carrier sentence than is often used in similar speech perception tasks was used, and this contained words that varied according to the accent of the carrier sentence.

The carrier sentence was produced in the two different accents by the same male speaker. This speaker was able to produce versions of both SSBE and Sheffield English that sounded like those of native speakers. This allowed direct comparison of the best exemplar locations chosen by subjects in both accents. Consequently, the task enabled the predictions about how a listener might adjust to a non-native accent based on the Perceptual Assimilation Model (Best, 1994) and the Speech Learning Model (Flege, 1992, 2003) to be tested. That is, by comparing the best exemplars that listeners chose in their native and non-native accents, it was possible to evaluate whether listeners were assimilating non-native vowels into their native accent categories, or whether they were creating a new phonetic category.

4. Study 1: The experiments[2]

4.1. Introduction

This study investigated whether listeners adjust their vowel categorization processes when listening to speech produced with different accents in the same language. The study contrasted two varieties of British English: SSBE and Sheffield English. Listeners with varying backgrounds were tested in their perception of these accents; northerners and southerners living in London (Experiment 1; Section 4.2) and northerners living in the north of England (Experiment 2; Section 4.3).

Listeners were tested using a vowel categorization task, the Goodness Optimization Task (Iverson and Evans, 2003; Evans and Iverson, 2004). The task uses an adaptive procedure based on computational algorithms to search within a high dimensional acoustic space for best exemplars of vowel categories. Listeners heard synthesized vowels embedded in natural carrier sentences that were produced in either a SSBE or Sheffield English accent. They gave goodness ratings on the vowels and a computer program iteratively adjusted F1, F2, F3, and duration values on successive trials. The aim was to assess whether listeners changed their best exemplar locations based on the accent of the carrier sentence. Of particular interest were the vowels in the experimental words *bath*, *bud*, and *cud*, that are produced differently in southern and northern English accents (Wells, 1982b; see Chapter 2).

---

[2] The material presented in this chapter was published in Evans and Iverson (2004).

4.2. Experiment 1: Northern and southern listeners living in London

4.2.1. Method

4.2.1.1. Subjects

Twenty-three subjects were tested. All subjects were paid for their participation. All were native, monolingual English speakers resident in London at the time of testing. They had lived in London for an average of 8.6 years and for a minimum of 1 year. The subjects were 20-45 years old, with an average age of 26.5 years, had no known hearing problems, and reported no speech, hearing or language difficulties.

Of the 23 subjects tested, 12 had a southern English accent background and 11 had a northern English accent background. Following the isogloss described by Chambers and Trudgill (1980), southern England is defined as the area to the south and east of Northampton, and northern England as the area extending north of Northampton to the Scottish border. As such, subjects originally from Scotland and Wales were not included in the test sample. The classification of background was based on where subjects had lived between the ages of 5 and 18 years, an important period for accent development. It is during this period that children's social identities, and concomitantly their regional accent, develop rapidly as they move from having a strong attachment to their caregiver to forming distinctive peer-groups during their adolescence (Kerswill and Williams, 2000; Williams and Kerswill, 1999).

4.2.1.2. Stimuli and Apparatus

The stimuli consisted of synthesized vowels in the phonetic environments /b/-V-/d/, /b/-V-/θ/, /k/-V-/d/, embedded in natural recordings of the carrier sentence *I'm asking you to say the word [ ] please.* There were 16 test words in total, covering the whole of the vowel space in both SSBE and Sheffield English: *bad*, *bard*, *bed*, *bird*, *bud*, *bod*, *bawd*, *bid*, *bead*, *booed*, *cud*, *could*, *cooed*, *birth*, and *bath*. *Bath* was included because northerners and southerners produce this word with different vowels; northerners say *bath* with the short vowel [a] whilst southerners produce this word with a longer back vowel, [ɑ]. *Beth* was included to control for consonantal environment effects. *Cud* and *could* were included in case the potential shift in the *bud* and *cud* vowels with accent was affected by lexical influences (i.e., if these words were produced with the northern [ʊ] vowel, *cud* and *could* would become homonyms, but *bud* would not become the same as any other lexical competitor). *Cooed* was included to better investigate fronting in the /u/ vowel. It is well-known that /u/ is undergoing fronting in both SSBE and northern English accents (e.g. Foulkes and Docherty, 1999) but it is unclear from the literature whether the amount of fronting is affected by consonantal context. As such, both a labial (/b/-V-/d/) and velar (/k/-V-/d/) contexts were included. As previously mentioned, the carrier sentence was specifically designed to give an overt cue to accent, as listeners were not told that they would be listening to either a northern or southern accent. The carrier sentence is longer than is typical for similar perceptual experiments and contained words that varied according to accent. In particular, the word *asking* was included in the carrier sentence because it exhibits the northern-southern [a ɑː] distinction; in SSBE *asking* is produced with a long, back vowel, [ɑː], but in SF it is produced with a short vowel, [a].

The carrier sentence was produced in both Sheffield English and SSBE accents by the same male speaker. The speaker was asked to speak at the same rate in each accent. As a result of differences in the phonetic realization of particular phonemes in each accent, the carrier sentence was shorter in length in Sheffield

English than SSBE: [amaskɪnjəʔseɪʔwɜːd---pliːz̥] in Sheffield English but [aɪmɑːskɪŋjuːtəseɪðəwɜːd---pliːz̥] in SSBE. Multiple instances of the carrier sentences were recorded in each accent and one from each accent was selected for use in the experiment. A spectrogram of each of the carrier sentences selected for use is given in Figs. 1 and 2. Vowel formant frequency measurements are given in Tables 6 and 7.

The speaker was born and brought up in Sheffield, having lived there until the age of 19 when he moved to the south of England where he had lived for 7 years at the time of recording. This speaker was selected because he was able to switch between accents at will, and was able to produce versions of both accents that were informally judged by trained phoneticians to sound like those of native speakers. He was coached to produce words such as *bath* with [ɑː] as despite using southern vowels in all other contexts, he normally produced these with the short vowel [a].

In addition to the carrier sentences, the speaker was recorded reading a 2-minute passage from a novel, *Harry Potter and the Philosopher's Stone* by J.K. Rowling (Appendix A), in both accents. The passage was selected because it contained a high number of words that are produced with a high back vowel, [ʊ], in northern English accents but with a low vowel, [ʌ] in southern English accents, and words that differ in their use of the vowels [a] and [ɑː] depending on whether they are produced in a southern or northern English accent.

CVCs were embedded in the carrier sentences. The bursts, fricatives, and aspiration were spliced from the sentence recording, and the voiced portions were synthesized online using the cascade branch of a Klatt synthesizer (Huckvale, 2003; Klatt and Klatt, 1990). This allowed for fine-grained coverage of the entire vowel space. Each stimulus had a middle portion in which the formant

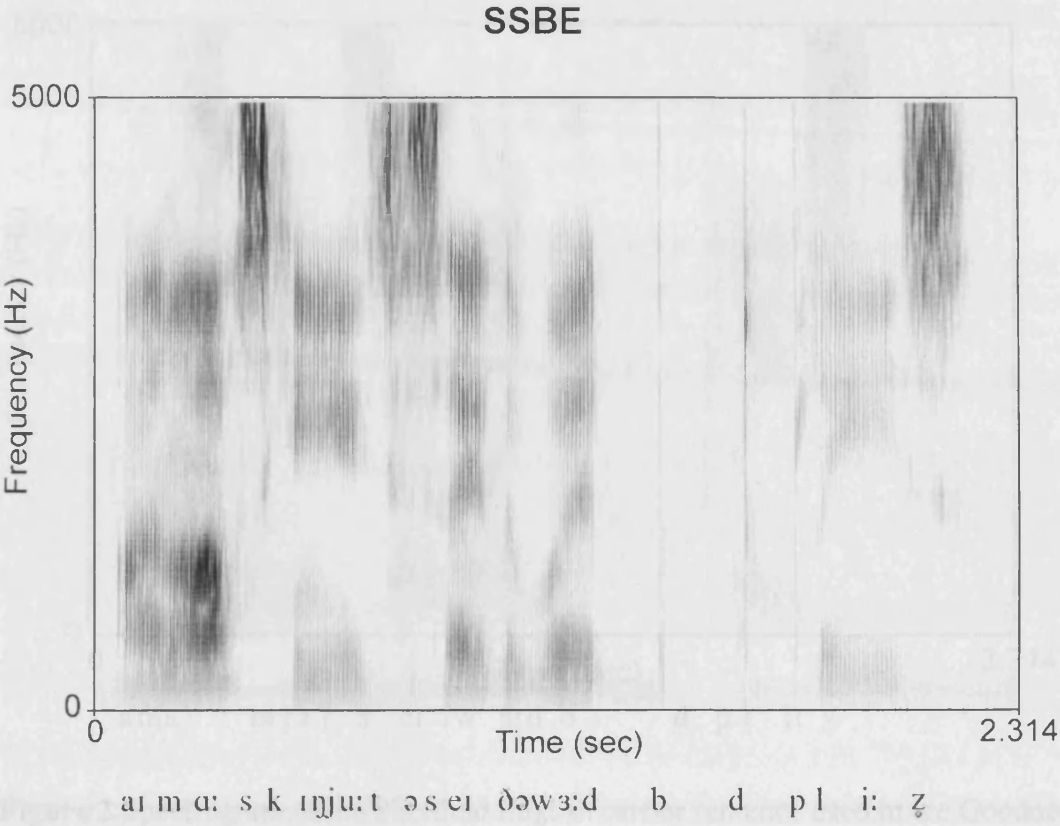

Figure 1 Spectrogram of the SSBE carrier sentence used in the Goodness Optimization Task.

Word	IPA transcription	F1	F2
I'm	ɑ	703	1161
	ɪ	614	1381
asking	ɑ	691	1138
asking	ɪ	371	2370
you	ʊ	299	2259
to	ə	467	1706
say	ɛ	498	1816
	ɪ	428	1972
the	ə	354	1569
word	ɜ	418	1701
please	ɪ	251	2296

Table VI F1 and F2 formant frequencies (Hz) for the vowels produced in the SSBE carrier sentence.

Figure 2 Spectrogram of the Sheffield English carrier sentence used in the Goodness Optimization Task

Word	IPA transcription	F1	F2
I'm	a	857	1332
asking	a	910	1325
asking	ɪ	371	2370
you	ə	418	1771
say	e	467	1719
	ɪ	424	1982
word	ɜ	467	1772
Please	i	315	2306

Table VII F1 and F2 formant frequencies (Hz) for the vowels produced in the Sheffield English carrier sentence.

frequencies were static, and had formant frequencies appropriate for the consonants (see below). All transitions were linear. The stimuli varied in terms of F1-F3 frequencies and duration of the middle portion. F1 frequency was restricted so that it had a lower limit of 150 Hz and an upper limit of 950 Hz. F2 was restricted to have a lower limit of F1+50 Hz, and had an upper limit defined by the equation;

$$F2_{upper\text{-}limit} = 3000\ Hz - 1.7*F1 \tag{1}$$

F3 frequency was restricted to have a lower limit of 2000 Hz, an upper limit of 3150 Hz and was always at least 100 Hz greater than F2. The duration of the middle portion was restricted to be greater than 20 ms and less than 403 ms.

All other synthesis parameters were chosen to mimic the natural speech recordings. For /b/-V-/d/, F1-F3 were 200, 1500, and 2400 Hz at the start of the formant transitions for /b/. The duration of the initial transition was 20 ms. F1-F3 were 200, 2300, and 3200 Hz at the end of the formant transitions for /d/. The final transition duration was 120 ms. F4 and F5 were fixed to 3200 and 4900 Hz respectively through the stimulus. The bandwidths of F1-F5 were fixed to 100, 120, 150, 100, and 175 Hz. F0 (fundamental frequency) started at 116 Hz, rose to 126 Hz, and fell to 104 Hz, following the speaker's natural production. AV (Amplitude of Voicing) started at 45 dB, rose to 51 dB, and fell to 48 dB.

For /k/-V-/d/, F1-F3 began at the target formant frequencies of the vowel. In other words there were no formant transitions before the onset of voicing. F1-F3 were 200, 1500, and 2600 Hz at the end of the formant transitions for /d/. The final transition duration was 40 ms. F4 and F5 were fixed to 3200 and 4450 Hz throughout the stimulus. The bandwidths of F1-F5 were fixed to 100, 120, 150, 150, and 175 Hz. F0 started at 125 Hz, rose to 128 Hz, and fell to 108 Hz. AV was increased from 0 to 40 dB over the first 10 ms, rose to 50 dB, and fell to 45dB.

For /b/-V-/θ/, F1-F3 were 200, 1300, and 2335 Hz at the beginning of the formant transitions for /b/. The duration of the initial transition was 20 ms. F2 and F3 were 1290 and 2400 Hz at the end of the formant transitions for [θ], and F1 ended on the target vowel frequency. The final transition duration was 20 ms. F4 and F5 were fixed to 3200 and 4900 Hz throughout the stimulus. The bandwidths of F1-F5 were fixed to 100, 160, 250, 150, and 175 Hz. F0 started at 115 Hz, rose to 125 Hz, and fell to 106 Hz. AV started at 40 dB, rose to 45 dB, and fell to 10 dB over the last 30 ms of the vowel.

After synthesis, the CVC stimuli were processed using a multi-band filter. This was to enable fine-tuning of the match between the long-term average spectra of the natural speech. Frequencies between 0 and 1500 Hz were attenuated by 1.5 dB, frequencies between 1500 and 3500 Hz were amplified by 6 dB, and frequencies between 3500 and 5500 Hz were attenuated by 2 dB. This filter was necessary because adjustments of the Klatt synthesis parameters, such as bandwidth, were not entirely sufficient by themselves to match the voice quality of the natural speech produced by the talker.

The stimuli were created on-line during the experiment and played at a sampling rate of 11 kHz using a computer sound card. Subjects listened over headphones (Sennheiser HD 414) in a sound attenuated booth.

4.2.1.3. Procedure

There were two testing sessions, one for each accent. The order of presentation was counterbalanced across subjects. Sessions were conducted on separate days to minimize the risk that subjects would be aware that the speaker was the same in both conditions. Subjects were informally questioned after completing the experiment and no subject reported that they had been aware that the speaker was the same. Each session was self-paced and lasted approximately 1 hour.

At the start of each session, subjects listened to the short passage read by the speaker in order to familiarize them with the accent. They then found the best exemplar for one practice word (*kid*), and best exemplars for the 16 test words: *bad, bard, bed, bird, bud, bod, bawd, bid, bead, booed, cud, could, cooed, Beth, birth,* and *bath.* To find the best exemplars, subjects heard a synthesized vowel embedded in a natural carrier sentence on a each trial, and were instructed to rate whether it was close to being a good exemplar of a target word that was displayed orthographically on a computer screen. The only instruction given to subjects was to rate whether the word that they heard sounded like a good example of that word to them. They were not given any explicit instruction to take the accent of the carrier sentence into account. They gave their response by positioning and clicking a computer mouse on a continuous scale from *close* to *far away*. This was shown on the screen as a bar. Subjects were able to listen to a trial again by clicking on a button marked 'Listen again'. The vowel parameters (F1, F2, F3, and duration) were adjusted after each trial using a customized computational procedure based on standard function minimization algorithms (Press et al., 1992). The procedure operated iteratively to find the best exemplar location for a particular word within this four-dimensional parameter space.

The procedure had five stages, with six trials per stage. It was able to find the best exemplar location within this large vowel space after 30 trials. Briefly, the procedure adjusted F1 and F2 in stages 1 and 2, starting along a path in Stage 1 that by passing through the average F1 and F2 frequencies that the speaker of the carrier sentence had used for that word, averaged over the two accents, would be likely to "home in" on a best exemplar most quickly. It then adjusted the more secondary dimensions of F3 and duration in Stages 3 and 4 and fine-tuned the best exemplar location in Stage 5.

The best exemplar was found in Stage 1 along a straight-line path through the F1/F2 plane that was defined by two points: the middle of the vowel space (F1 = 500 Hz, F2 = 1500 Hz) and the average F1 and F2 frequencies that the speaker of the carrier sentence had used for that word, averaged over the two accents. The path passed through these points and ended at the boundaries of the vowel space. For example, the Stage 1 search path for *bead* crossed diagonally through the vowel space, from the extreme high-front boundary of the space (i.e., high F1, low F2 near /i/), through the speaker's measured values for /i/ and the middle of the vowel space, and through to the extreme low-back boundary of the space (i.e., low F1, high F2, near /A/). All other parameters were fixed at neutral values for this stage; F3 was set at 2500 Hz and duration at 116 ms.

On the first two trials of Stage 1, subjects heard the most extreme stimuli that it was possible to synthesize along the search path. For example, for *bead*, subjects heard extreme high-front and low-back vowels. The order of these two trials was randomized. For the remaining stimuli, the selection of stimuli was based on listeners' goodness ratings. Formulae were used that were designed to find stimuli along the path that would be perceived by the listener to be better exemplars. On the third trial, subjects heard a stimulus that was selected using a weighted average of the ratings for the first two stimuli, according to the equation;

$$c = a * \frac{f(b)}{f(a) + f(b)} + b * \frac{f(a)}{f(a) + f(b)} \qquad (2)$$

where *a* and *b* are the positions on the search path for the first two trials, *f(a)* and *f(b)* are the goodness ratings for the stimuli on those trials (the goodness ratings were scaled from 0 to 1), and c is the new path position selected for the 3rd trial. On the 4^{th}, 5^{th} and 6th trials, the stimuli were selected by finding the minimum of a parabola that was defined by the equation;

$$\min = \frac{b - 0.5 * \{[b-a]^2 * [f(b) - f(c)] - [b-c]^2 * [f(b) - f(a)]\}}{[b-a] * [f(b) - f(c)] - [b-c] * [f(b) - f(a)]} \qquad (3)$$

where b was the path position of the best stimulus found thus far; a and c were the most recently tested positions on either side of b; and $f(a)$, $f(b)$, and $f(c)$ were the goodness ratings for those stimuli. In cases where Eq. (3) could not be calculated (e.g. if a, b, and c were co-linear, or b was at an extreme position on the path), a weighted average [Eq. (2)] was calculated instead, based on the parameters of the best stimulus found thus far and the last stimulus that had been played. At the completion of this stage, the parameters of the best stimulus found thus were passed onto the next stage of the search algorithm.

The same six-stage search algorithm was used for the other stages, but along different paths. Stage 2 found the best exemplar along a straight-line path that was orthogonal to the Stage 1 path in the F1/F2 plane, and included the best exemplar found in Stage 1. Stage 3 searched along the F3 dimension, keeping all parameters fixed to the best exemplar values that had been found at the end of Stage 2. Stage 4 searched along the duration dimension, keeping all other parameters fixed to the best exemplar values found in Stage 3. Duration was scaled using log values. This was considered to be a closer representation of listeners' sensitivity to differences in duration than a linear scale. Stage 5 searched along a straight-line path through a three-dimensional F1, F2, and duration space (F3 did not vary) that began in the middle of the vowel space (F1 = 500 Hz, F2 = 1500 Hz, and duration = 116 ms) and passed through the parameters of the best exemplar chosen thus far.

At the end of each stage, listeners heard the best exemplar found thus far, and responded whether they thought it was close to being a good exemplar of that category. Subjects were allowed to repeat stages if they responded that the best exemplar found thus far was poor. This typically happened when the search algorithm got thrown off by an erroneous goodness rating. The best exemplar was defined as the stimulus given the highest goodness rating in Stage 5.

4.2.2. Results

Three subjects were dropped from the experiment because their best exemplar locations were unreliable. Reliability was established by examining subjects' best exemplars for the words *bird*, *bed*, *bead*, and *bad* in both SSBE and Sheffield English carrier sentences. These words were chosen as they are produced in the same way in both northern and southern English accents. Subjects who chose best exemplars for these words that differed by more than 2 ERB (Equivalent Rectangular Bandwidth: see Moore et al., 1997) between the two carrier sentences were excluded from the study. Of the remaining 20 subjects, 10 were from southern England and 10 were from northern England.

4.2.2.1. Bud and Cud

As displayed in Fig. 3, listeners chose different formant frequencies for *bud* and *cud* in SSBE and Sheffield English carrier sentences. The shift appeared to occur predominantly in the F1 dimension; both groups of listeners chose a higher F1 for *bud* and *cud* in SSBE carrier sentences than in Sheffield English sentences, although the size of the shift appeared to be larger for northerners.

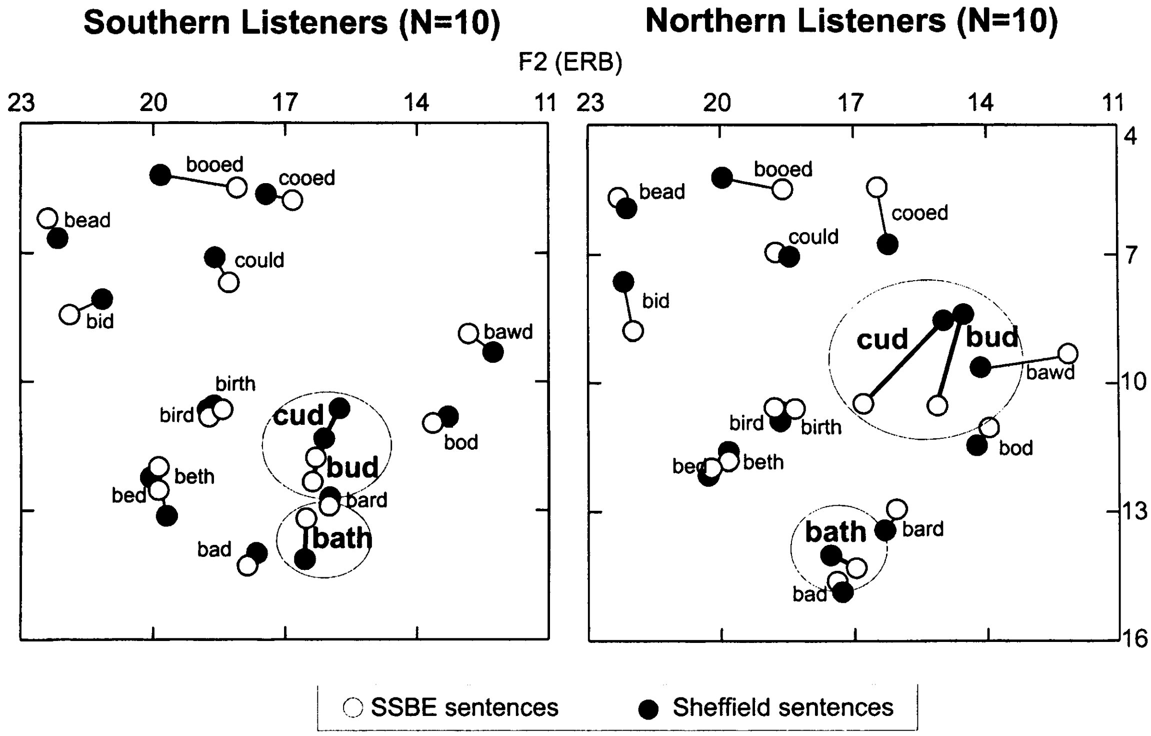

Figure 3 Average F1 and F2 formant frequencies (ERB) of best exemplars for northern and southern listeners in SSBE and Sheffield English carrier sentences. The F1 frequencies of *bud* and *cud* were significantly different in the two carrier sentences for both groups of listeners, but no other words were reliably normalized for accent.

The differences in F1 and F2 were tested in separate repeated measures ANOVA (Analysis of Variance) analyses, with word (*bud* or *cud*) and sentence context (SSBE or Sheffield English) coded as within-subject variables, and subject background (northern or southern) coded as a between-subject variable.

For F2, there were no significant main effects or interactions, $p > 0.05$, suggesting that listeners were not normalizing for accent in this dimension. For F1, however, there was a main effect of sentence context, $F(1,18) = 11.94, p < 0.01$, confirming the observation that all listeners chose higher F1 frequencies for *bud* and *cud* in the SSBE sentences. There was also a main effect of subject background, $F(1,18) = 12.08$, $p < 0.01$, demonstrating that northern listeners consistently chose higher F1 values for *bud* and *cud* than southern listeners. There was no significant main effect of word, $p > 0.05$, demonstrating that normalization was not affected by lexical status, and no significant interactions, $p > 0.05$, indicating that the observation that northerners were normalizing more than southerners was not reliable.

The effects of sentence context and subject background on F1 can be seen clearly in Fig. 4. In the Sheffield English sentences, northerners chose a high back vowel (i.e. low F1 frequency) that was appropriate for that accent, but southerners chose a low-central vowel with a higher F1 frequency than Sheffield English speakers actually produce. In the SSBE context, southerners chose a low vowel (i.e. high F1 frequency) that was appropriate for that accent, but northerners chose a central vowel that is higher (i.e. lower F1 frequency) than SSBE speakers actually produce. Although the size of the shift in *bud* was relatively small for southerners, the direction of this shift was consistent; nine out of ten southerners chose lower F1 frequencies for *bud* in the Sheffield context.

As displayed in Tables VIII and IX, there were few differences between best exemplars of *bud* and *cud* in terms of F3 or duration. Separate repeated measures ANOVA analyses for F3 and duration were used to confirm these observations. The analyses demonstrated that there were no significant main effects or interactions of sentence context, subject background, or word, $p > 0.05$, for F3 or duration. This further confirms that vowel normalization for accent only took place in the F1 dimension for *bud* and *cud*.

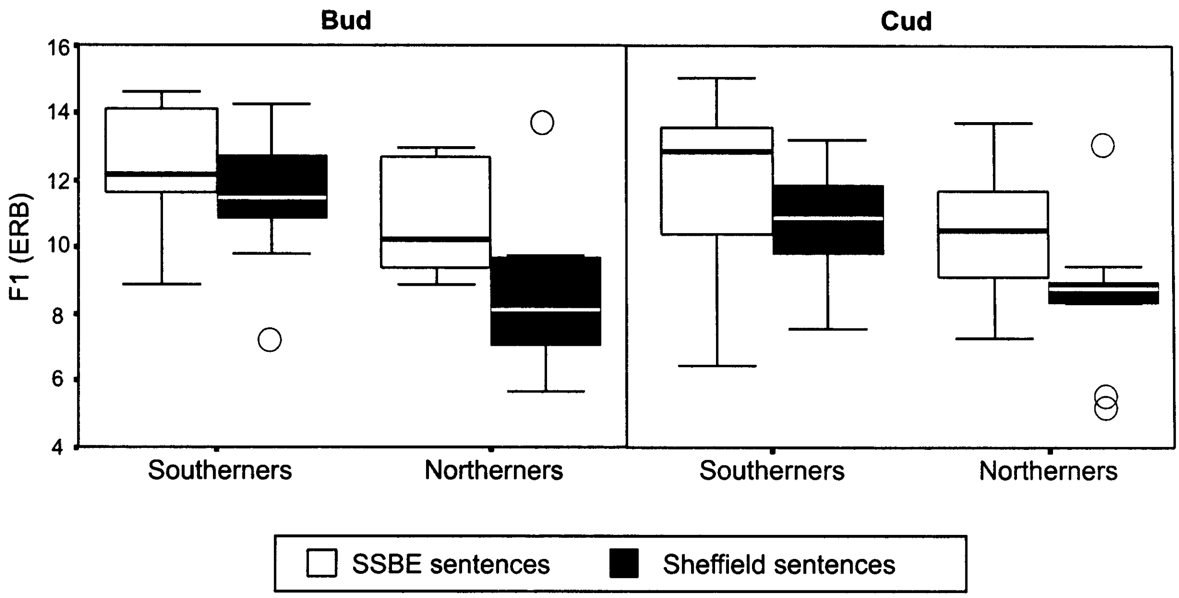

Figure 4 Boxplots of F1 formant frequency values (ERB) for *bud* and *cud* in SSBE and Sheffield English carrier sentences for northern and southern listeners. The boxplots display the interquartile range of scores. The box shows the 25th to 75th percentiles, with a line at the median value. The lower and upper "whiskers" respectively show the first and last quartiles, with outliers represented by the unshaded circles. The best exemplar locations for northerners had lower F1 frequencies than those chosen by southerners, and both groups of listeners chose lower F1 frequencies in Sheffield than in SSBE carrier sentences.

Table VIII Average F3 frequencies (ERB) of best exemplars for northern and southern listeners in SSBE and Sheffield English sentence contexts.

	Northern		Southern		
Word	**SSBE**	**Sheffield**	**SSBE**	**Sheffield**	**Average**
bud	23.4	23.0	22.8	23.5	23.2
cud	23.7	23.2	22.9	23.4	23.3
bath	22.7	22.1	23.1	22.1	22.5
bead	24.0	24.5	24.0	23.7	24.1
bid	24.0	24.2	23.8	23.2	23.8
bed	22.9	22.9	23.2	23.2	23.1
beth	23.4	22.4	23.2	23.4	23.1
bird	22.7	22.4	22.9	23.0	22.8
birth	22.3	22.4	23.5	23.1	22.8
bad	22.1	22.1	22.6	22.9	22.4
bard	22.8	22.9	23.6	23.2	23.1
bod	23.6	23.3	23.1	22.7	23.2
bawd	23.6	23.2	23.5	23.3	23.4
booed	22.2	22.4	22.4	22.4	22.4
cooed	22.5	22.8	22.2	23.4	22.7
could	23.3.	23.0	23.4	22.8	23.1
Average	23.1	22.9	23.1	23.1	

Table IX Average durations (ms) of best exemplars for northern and southern listeners in SSBE and Sheffield English sentence contexts.

	Northern		Southern		
Word	**SSBE**	**Sheffield**	**SSBE**	**Sheffield**	**Average**
bud	62.5	72.8	74.5	62.2	68.0
cud	63.7	62.0	81.8	69.7	69.3
bath	81.8	71.5	146.0	108.6	102.0
bead	138.9	129.4	120.9	130.7	130.0
bid	72.5	60.0	63.0	62.5	64.5
bed	78.1	66.8	78.1	66.8	72.5
beth	76.9	68.1	71.3	71.6	72.0
bird	182.6	144.2	119.5	130.6	144.2
birth	135.6	137.5	145.1	158.1	144.1
bad	101.0	70.6	104.6	84.7	90.2
bard	208.2	181.3	176.6	174.0	185.0
bod	68.5	79.5	77.5	74.4	75.0
bawd	187.8	170.1	176.3	151.7	171.5
booed	176.7	161.7	162.1	154.5	163.8
cooed	173.3	169.7	147.0	146.0	159.0
could	65.3	64.8	70.0	77.2	69.3
Average	117.1	106.9	113.4	107.7	

4.2.2.2. Bath

As displayed in Fig. 3, listeners chose relatively similar formant frequencies for *bath* in SSBE and Sheffield English carrier sentences. There was perhaps a small shift in the F1 dimension for southern listeners. Separate repeated measures ANOVA analyses for F1, F2, and F3 demonstrated that there were no significant main effects or interactions of sentence context or subject background, $p > 0.05$, suggesting that listeners did not change the formant frequencies of *bath* to match the accent of the carrier sentence.

As displayed in Table IX, there were no strong normalization effects for duration; listeners chose similar vowel durations in both sentence contexts, although there was a trend for southerners to choose shorter vowels in the Sheffield English sentences. However, there appears to be a consistent effect of subject background; southern listeners chose a longer vowel for *bath* in both sentence contexts than northerners. A repeated measures ANOVA analysis verified that there was a main effect of subject background, $F(1,18) = 8.09, p < 0.01$, and that there was no significant main effect of sentence context or significant interactions, $p > 0.05$. The effect of subject background on duration can be clearly seen in Fig. 5. Northerners preferred a shorter vowel that corresponded to their production of [a] in *bath*, whereas southerners preferred longer vowels that corresponded to their production of [ɑː] in *bath*.

Although listeners did not choose significantly different formant in the SSBE and Sheffield English contexts, the results trended in the same direction (see Fig.3). The median values for F1 and F2 in both SSBE and Sheffield English sentence contexts were more similar to *bad* ([bad]) than *bard* ([bɑːd]) for northerners, and were more similar to *bard* than *bad* for southerners. This difference may have failed to reach significance because [a] and [ɑː] have very similar formant frequencies; the vowels differ more markedly in duration.

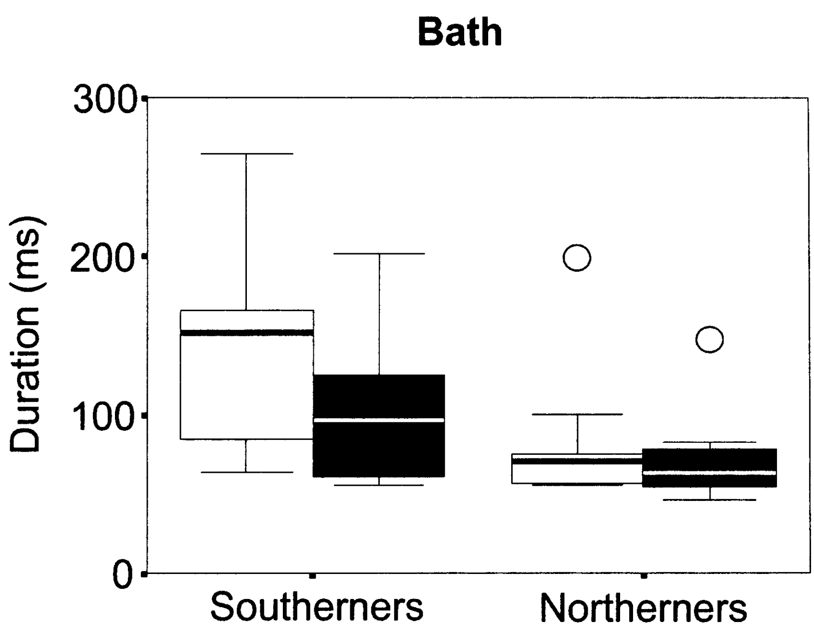

Figure 5 Boxplots of duration values for bath in SSBE and Sheffield English carrier sentences for northern and southern listeners. Northerners chose shorter vowels than southerners overall, but there was no normalization for the accent of the carrier sentences.

4.2.2.3. Other words

As displayed in Fig. 3, listeners chose similar F1 and F2 frequencies in SSBE and Sheffield English carrier sentences for most other words, although there were shifts for *bawd* and *booed*. The potential differences in F1 and F2 were tested in separate repeated measures ANOVA analyses, with word (i.e., all words other than *bud*, *cud*, and *bath*) and sentence context coded as within-subject variables, and subject background as a between-subject variable. There was a main effect of word for F1, $F(12,216) = 203.98, p < 0.01$, and F2, $F(12,216) = 113.76, p < 0.01$, demonstrating that listeners chose different formant frequencies for different words, but there were no main effects of sentence context or subject background, and no significant interactions, $p > 0.05$. The differences in *bawd* and *booed* shown in Fig. 3 were thus not reliable.

As displayed in Tables VIII and IX, listeners generally chose similar values for F3 and duration. Separate repeated measures ANOVA analyses revealed that there was a main effect of word for F3, $F(12,216) = 7.78$ $p < 0.01$, and for duration, $F(12,216) = 40.58$, $p < 0.01$, demonstrating that listeners chose different F3 and duration values for different words, but there were no main effects of sentence context or subject background, and no significant interactions, $p > 0.05$.

It is notable that both northern and southern listeners chose a high-front vowel for *booed* and a high-central vowel for *cooed* (see Fig. 3), rather than high-back vowels with lower F2 frequencies, as might have been expected based on earlier descriptions of British English (e.g. Wells, 1982b). Although these preferences may seem unusual, they correspond to recent changes in the way that British English speakers produce these vowels. Younger speakers in particular have begun to produce these vowels with less lip rounding and a more forward tongue position (Docherty and Foulkes, 1999; Kerswill and Williams, 2000; Torgersen, 1997; Williams and Kerswill, 1999).

4.3. Experiment 2: Ashby listeners

4.3.1. Introduction

Experiment 1 demonstrated that listeners living in London normalized *bud* and *cud* according to the accent of the carrier sentence, and that the patterns of normalization depended on whether the listeners were from a northern or southern background. The northerners tested in Experiment 1 had been living in the south of England for a minimum of 1 year, and had experience of interacting with SSBE speakers. It is possible that this experience may have affected their vowel categorization processes. Experiment 2 examined the role of language experience on vowel normalization by testing northerners who were still living in the north of England. Subjects were born and raised in Ashby de la Zouch, a small market-town in North-West Leicestershire, where the dominant accent is similar to that spoken in Sheffield. Like the regional accent spoken in Sheffield, the accent spoken in Ashby is a northern variety of English. Thus, words like *bath* are produced with a short vowel [a] (e.g., [baθ]), and words like *bud* and *cud* are produced with a high-back vowel, [ʊ], such that *cud* and *could* are homophones or near-homophones.

Subjects were aged 16-17 years old, and had not yet moved for employment or university education. Unlike large cities such as London, Ashby is not a multidialectal environment where other regional accents are regularly used. Although the town has recently grown in population as a result of the building of a major route connecting it with Nottingham to the north and Birmingham to the south-west, the majority of the population is local to the area. In particular, it is highly unusual to find native speakers of SSBE who have moved to the town. Although people living in Ashby are highly familiar with SSBE through the media (Foulkes and Docherty, 1999), unlike subjects in London, they have not had experience of interacting with speakers of a wide variety of regional accents. The aim of the experiment was to determine whether the patterns of normalization found for northerners in Experiment 1 were affected by subjects' experience of living in London, or whether all northerners (i.e., even those who have not lived in the south of England) have the same patterns of normalization.

4.3.2. Method

4.3.2.1. Participants

Twelve subjects were tested and were paid for their participation. All were native monolingual English speakers, aged 16-17 years, who had been born and raised in Ashby. No subject reported any speech, hearing or language problems. As in Experiment 1, subjects were dropped from the experiment if their best exemplar locations for four vowels that are produced in the same way in northern and southern English accents (*bead*, *bed*, *bad* and *bird*) differed by more than 2 ERB. One subject was dropped from the experiment, because her best exemplar locations for these words were unreliable.

4.3.2.2. Stimuli and Apparatus

The stimuli were synthesized in advance so that the experiment could be run using a laptop computer. The entire range of possible vowels was synthesized with a resolution of 0.5 ERB in F1 and F2. Even though the stimulus space was quantized using ERB values, the algorithm searched through the vowel space along linear Hz paths, following the methodology used in Experiment 1. The linear Hz values were rounded in the search algorithm to the values of the nearest available stimulus. This allowed the search algorithm to be the same as in Experiment 1, and still have a stimulus quantization that was perceptually uniform. Duration was quantized in 16 steps on a log scale, from 20 to 403 ms. F3 was fixed at 2500 Hz for all stimuli; although Experiment 1 had shown that F3 varied for different words, the results suggested that this parameter made only a modest contribution to perceived goodness. There were a total of 7616 stimuli synthesized for each of the CVC contexts. The stimuli and the apparatus were the same as in Experiment 1 in all other respects.

4.3.2.3. Procedure

There was a four-stage search for best exemplars along the F1, F2, and duration dimensions, with six trials for each stage. The F3 adjustment stage was omitted. The procedure was the same as Experiment 1 in all other respects.

4.3.3. Results

4.3.3.1. Bud and Cud

As displayed in Fig. 6, Ashby listeners chose similar formant frequencies for *cud* in SSBE and Sheffield carrier sentences, but for *bud* there was a possible shift in the F2 dimension; listeners tended to choose a higher F2 for *bud* in SSBE than in Sheffield carrier sentences. As in Experiment 1, the potential differences in the formant frequencies were tested in separate repeated measures ANOVA analyses for F1 and F2 with sentence context (SSBE or Sheffield English) and word (*bud* or *cud*) coded as between-subject variables, and subject background as a between-subject variable. For F1 and F2, there were no main effects of word, sentence context, or subject background, $p > 0.05$, demonstrating that the shift in the F2 dimension for *bud* was not reliable.

As displayed in Table X, there was also little difference between *bud* and *cud* in terms of duration. A repeated measures ANOVA analysis revealed that there were no main effects of word, sentence context, or their interactions, $p > 0.05$.

There was thus no consistent evidence to suggest that Ashby listeners normalized *bud* and *cud* for accent. Instead, they chose traditionally northern vowels for *bud* and *cud* in both northern and southern English carrier sentences.

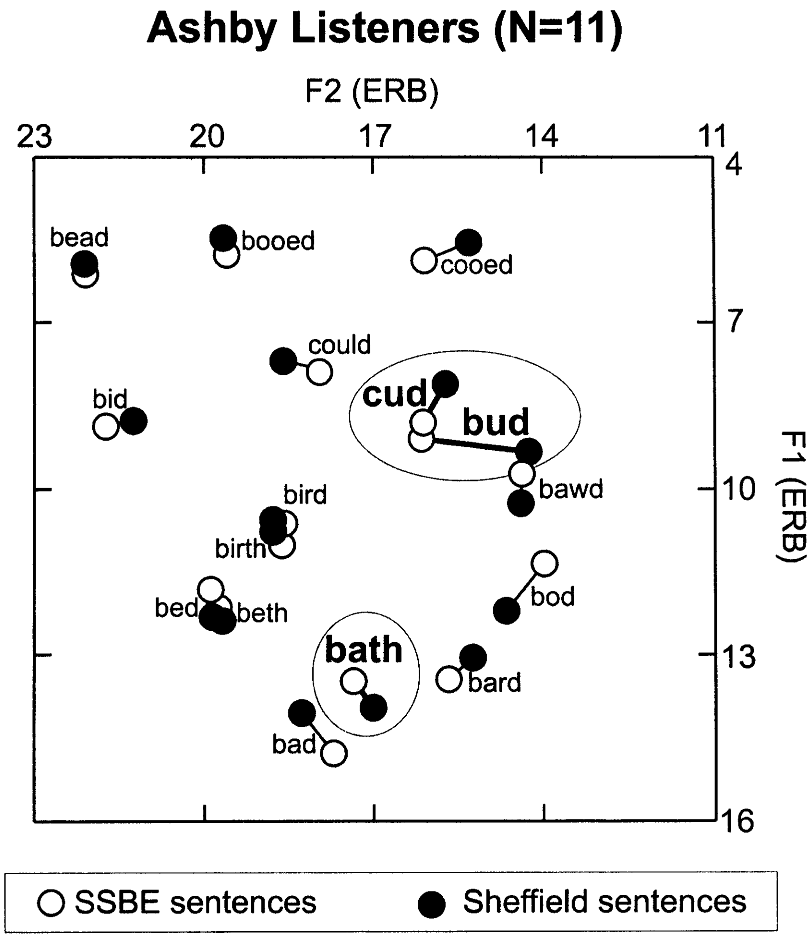

Figure 6 Average F1 and F2 formant frequencies (ERB) of best exemplar locations for Ashby listeners in SSBE and Sheffield English carrier sentences. The F1 and F2 frequencies did not vary significantly between the two carrier sentences, suggesting that no vowels were normalized for accent.

Table X Average durations (ms) of best exemplars for Ashby listeners in SSBE and Sheffield English sentence contexts.

Word	SSBE	Sheffield	Average
bud	63.8	65.4	64.6
cud	76.9	55.8	66.4
bath	108.4	79.3	93.8
bead	99.0	115.7	107.4
bed	84.5	56.5	70.5
bid	71.9	68.5	70.2
beth	73.9	58.2	66.0
bird	139.9	117.3	128.6
birth	124.5	113.3	118.9
bad	85.4	77.8	81.6
bard	156.4	162.7	159.5
bod	72.4	72.0	72.2
bawd	140.1	124.5	132.3
booed	170.4	164.4	167.4
cooed	146.7	140.8	143.8
could	76.1	69.6	72.9
Average	105.6	96.4	

4.3.3.2. Bath

Like the northern listeners tested in Experiment 1, Ashby listeners also chose similar formant frequencies for *bath* in SSBE and Sheffield carrier sentences (see Fig. 6). There was a trend for listeners to choose a longer vowel in the SSBE than in the Sheffield English carrier sentences ($F(1,10)=3.52$, $p < 0.1$; see Table X). However, separate repeated measures ANOVA analyses for F1, F2, and duration revealed that there was no main effect of sentence context, $p > 0.05$. There was thus no clear evidence to suggest that Ashby listeners normalized *bath* for accent.

4.3.3.3. Other words

As in Experiment 1, listeners chose similar formant frequencies for all other target words in SSBE and Sheffield English carrier sentences (see Fig. 6). Listeners also chose similar durations for each target word in both sentence contexts (Table X). Separate repeated measures ANOVA analyses for F1, F2 and duration revealed that there was a main effect of word for F1, $F(12,120) = 76.08$, $p < 0.01$, F2, $F(12,120) = 47.24$, $p < 0.01$, and duration, $F(12,120) = 15.61$, $p < 0.01$. However, there was no significant interaction of F1, F2 or duration with word, $p > 0.05$, suggesting that none of the other words varied according to the accent of the carrier sentence.

4.4. General Discussion

The results of Study 1 demonstrated that individuals living in London normalized the vowels in *bud* and *cud* – but not *bath* – for southern and northern English accents, and that the patterns of normalization reflected each listener's linguistic experience. When individuals living in London heard sentences that were similar to their native accent, they chose formant frequencies for *bud* and *cud* that matched what speakers of that accent would produce; they chose a long, low-back vowel ([ɑː] for *bath*, and a low-central vowel [ʌ] for *bud* and *cud*. Southerners living in London preferred the low-central vowel [ʌ] when listening to SSBE sentences, and northerners living in London preferred the high-back vowel [ʊ] when listening to Sheffield English sentences. When individuals living in London heard sentences that did not match their native accent (i.e., northerners listening to SSBE and southerners listening to Sheffield English sentences), they chose centralized vowels for *bud* and *cud* instead of the [ʌ] and [ʊ] vowels which would normally be produced in SSBE and Sheffield English accents, respectively. Northerners who were less experienced with southern accents (i.e., Ashby listeners) did not normalize for accent at all. Instead, they chose vowels in SSBE and Sheffield English sentences that match what northern speakers would produce.

Episodic memory research has shown that individuals store phonetically detailed representations of spoken words in long-term memory (e.g., Goldinger, 1996, 1998; Nygaard and Pisoni, 1998; Palmeri et al., 1993). Based on this research, it was hypothesized that listeners would choose best exemplars based on their long-term memory representations for words produced by speakers with similar accents. It was surprising then that northerners living in London, who have a lot of experience of listening to and interacting with SSBE speakers, chose best exemplars for *bud* and *cud* in SSBE sentences that do not match how SSBE speakers produce these vowels; northern listeners preferred a central vowel for *bud* and *cud* in SSBE carrier sentences, but southerners produce these words using a lower vowel, [ʌ]. This suggests that listeners were not performing the task based on long-term memory representations that accurately reflect how SSBE speakers talk.

Although this may suggest that listeners were not performing the task using stored exemplars, it is possible that listeners were using inaccurate exemplars that had been affected by perceptual magnet effects (Iverson and Kuhl, 1995, 1996, 2000;

Iverson et al., 2003; Kuhl, 2000) or category assimilation processes (Best et al., 1988, 2001; Flege, 1992, 1995). That is, northerners' perception of SSBE [ʌ] may have been distorted because they do not have a native /ʌ/ category; northerners may perceive the SSBE [ʌ] to be a member of their native /ə/ or /ɜː/ categories [or a new category caused by merging /ʌ/-/ʊ/ (Flege, 1995; MacKay et al., 2001)], causing the SSBE [ʌ] to sound more centralized. It is plausible that such perceptual distortion caused northern listeners to remember mistakenly that southerners produce centralized vowels for *bud* and *cud*. In other words, the basic hypothesis that listeners choose best exemplars that match long-term memory representations may be correct, but the memories of listeners may be inaccurate.

There are two aspects of the present results that are inconsistent with this perceptually distorted exemplar account. First, the *bud* and *cud* vowels that southerners chose in the Sheffield sentences cannot be easily explained by perceptual magnet effects or category assimilation. Northerners produce *bud* and *cud* using the vowel [ʊ]. Southerners already have this vowel category, although they use it with a different lexical distribution. Thus, one would expect southerners to 'assimilate' the northern [ʊ] vowel in *bud* and *cud* into their own native /ʊ/ category. Instead, southerners chose a low-central best exemplar for *bud* and *cud* that is on the other side of the vowel space to [ʊ]. It seems unlikely that southerners erroneously perceive the northern [ʊ] as a low-central vowel. Second, there was no normalization for *bath*. Southerners and northerners both use the vowels [a] and [ɑː] but with different lexical distributions, and speakers of British English are very aware that this difference is a clear marker of accent (Trudgill, 1986; Wells, 1982b). Thus, northerners in London know that southerners produce *bath* with a long vowel [ɑː], and southerners in London know that northerners produce *bath* with a short vowel [a]. Yet listeners in this experiment chose to base their choice of exemplar on their own accent, rather than on their knowledge of what a speaker of that accent would produce.

Although these patterns of normalization may appear idiosyncratic, they correspond closely with the changes in production that speakers tend to make as a result of sociolinguistic factors when they live in multidialectal environments (Trudgill, 1986). Northerners who live in the south of England typically modify aspects of their accent to fit in with southerners. Phonetic descriptions of northerners accommodating to southerners have shown that northerners modify the vowel they use in words like *bud* and *butter* (Trudgill, 1986; Wells, 1982). When northerners change their accent so that it is closer to SSBE, they produce these words using a vowel intermediate between the SSBE STRUT vowel ([ʌ]) and the FOOT vowel ([ʊ]) used in both SSBE and Sheffield English. Such intermediate qualities include a mid-back [ʊɠ]; the unrounded equivalent, [ɣ]; a half–open vocoid, unrounded or slightly rounded, similar to cardinal [ʌ] (and therefore somewhat different from RP /ʌ/ that is usually central rather than back), and a mid or half-close [ə] (see Chapter 2 for a more detailed description).These vowels are much like the vowel that northerners chose as best exemplars for these words in SSBE carrier sentences.

However, northerners do not change all aspects of their production to match SSBE speakers. Although many northerners change the way in which they produce words like *bud* and *cud* to better match how southerners produce these words, the short vowel used in words like *bath* is often retained. As Wells (1982b: 354) describes, "there are many educated northerners who would not be caught dead doing something so vulgar as to pronounce STRUT words with [ʊ], but who would feel it to be a denial of their identity as northerners to say BATH words with anything other than short [a]" (see also Trudgill, 1986: 18). There thus appears to be a tension between the desire to fit in with a new community, and a concurrent wish to continue signalling aspects of their own regional identity through their speech. In the same way that northerners retain [a] when producing words like *bath* then, the subjects in this study continued to choose short [a] for *bath* in both SSBE and Sheffield English carrier sentences.

Southerners living in London are less inclined to modify their accent when speaking to others though, because theirs is the dominant accent of the community. Likewise, they made little adjustment to *bud* and *cud*, and they preferred a southern pronunciation of *bath* in both SSBE and Sheffield English carrier sentences.

Production may also help to explain why Ashby listeners did not perceptually normalize for accent. Differences in the amount of experience a listener has with an accent within the same language have been shown to affect perceptual processing. For example, listeners who are highly experienced with a particular accent have been shown to perform better in word recognition tasks than those with little experience, particularly when identifying words that are produced with markedly different vowels (Labov and Ash, 1997). One could imagine then, that Ashby listeners did not normalize because they had not yet had enough perceptual experience with southern accents. However, Ashby listeners are highly familiar with southern British English accents. They are regularly exposed to SSBE speakers through the media (Foulkes and Docherty, 1999) and are able to identify a speaker of SSBE correctly in perceptual tests (Evans, 2001). Moreover, all listeners heard a short passage read by the speaker in the relevant accent before starting the experiment, and such short-term exposure has been shown to be enough to tune speech recognition processes to the characteristics of individual talkers (Nygaard and Pisoni, 1998).

It may have been more important that these subjects (aged 16-17 years, born and raised in Ashby) had no experience of modifying their own speech in order to fit into a multidialectal community (e.g., when attending university). It is possible that experience gained through interaction with a native speaker provides cues for learning that are not present in the raw auditory signal. Indeed, social interaction has been shown to improve discrimination of a non-native phonetic contrast in young infants. Kuhl et al. (2003) tested two groups of 9-month-old American infants in their perception of Mandarin Chinese. One group of infants was exposed to Mandarin Chinese through interaction with a speaker. Another group was exposed to the same foreign-language speakers and materials but through audio-visual or audio-only recordings. Discrimination of Mandarin contrasts improved in infants who were exposed to Mandarin through interaction with a native speaker, but infants who were only exposed to Mandarin through audio-only or AV recordings did not change in their performance. It is plausible then, that cues gained through social interaction remain important in phonetic learning at a later stage in development, and that these Ashby listeners chose northern vowels in SSBE carrier sentences because they had not yet had experience of interacting with southerners, even though they know how southerners talk.

5. Study 2: A longitudinal study of accent change in young adults

5.1. Introduction

Study 1 demonstrated that some subjects adjusted their vowel categorization decisions when listening to speech produced in different regional accents. Subjects living in London normalized the vowel in *bud* and *cud* – but not *bath* – for southern and northern English carrier sentences, with the patterns of normalization reflecting each listener's linguistic experience.

When individuals living in London heard sentences that were similar to their native accent, they chose formant frequencies for *bud* and *cud* that matched what speakers of that accent would produce. When individuals living in London heard sentences that did not match their native accent (e.g., northerners listening to SSBE speech), they chose centralized vowels for *bud* and *cud* rather than the [ø] and [U] vowels that would normally be produced in SSBE and Sheffield accents respectively. These patterns of normalization corresponded closely with the changes in production that speakers tend to make when they live in a multidialectal environment (Wells, 1982b; Trudgill, 1986). Northerners who live in the south of England typically modify some aspects of their accent in order to fit in with southerners; they change their production of the vowel in words like *bud* and *cud* so that it becomes more centralized. Northerners also retain some aspects of their regional accent though; they retain [a] when producing words like *bath* (Wells, 1982b; Trudgill, 1986; see Chapter 2).

However, not all subjects changed their vowel categorization decisions according to the accent of the carrier sentence. Northerners who were less experienced with southern accents (i.e., Ashby listeners) did not normalize for accent at all. Instead, they chose vowels in SSBE and Sheffield English that match what northern speakers would produce.

It was hypothesized that production might be able to explain why Ashby listeners did not perceptually normalize for accent. Although these listeners were highly familiar with SSBE and regularly exposed to SSBE speakers through the media (Foulkes and Docherty, 1999), they had not had experience of modifying their own speech in order to fit into a multidialectal community. Consequently, it was hypothesized that experience with a non-native accent, gained through interaction with a native speaker might provide cues for learning that are not present in the raw auditory signal (see also Kuhl et al., 2003). The aim of this study then, was to investigate the role of social interaction in the ability to adapt to a non-native accent. The study aimed to determine whether information gained through experience of interacting with speakers of a non-native accent caused subjects to change their own production, and whether these changes also affected their perceptual processes (i.e., their best exemplar locations for vowels produced in a non-native accent, and their ability to recognize the accent in noise).

5.2. The relationship between production and perception: Do changes in production affect perception?

There are a number of reasons to think that there should be a close link between speech production and perception, and that changes in one domain might affect the other. We learn to perceive and produce our native language as infants by listening to what other speakers produce, and in some way we come to associate the sounds we hear with the way of producing them (see e.g. Kuhl and Meltzoff, 1996). Learning to produce the sounds of our native language also depends on hearing oneself produce sound (Kuhl, 2000; Vihman and Nakai, 2003). Since we are likely to have heard our own productions more than any other single individual, it is likely that our own production, and therefore our native language and probably our native accent, will have played a particularly important role in the development of our perceptual system.

Evidence to support a link between production and perception comes from studies of L2 acquisition that have found that experience in either the perception or production of a non-native accent or language has affected the other. Bradlow et al. (1997) demonstrated that adult Japanese speakers who were trained in the perception of /r/ and /l/ showed significant improvements in their production; native English listeners were able to identify their productions more accurately after they had completed the perceptual training than before. Meador et al. (1997) showed that sentence comprehension affected the ability to produce a foreign language accurately. Native Italian speakers living in Canada were tested in their comprehension and subsequent production of semantically unpredictable sentences. Participants heard a sentence and then repeated it back as accurately as possible. Native-English speakers then rated participants' production of the sentences for degree of foreign accent. The results demonstrated that there was a relationship between the foreign accent ratings and the number of words the Italian subjects were able to repeat; the larger the number of words they were able to repeat, the lower their foreign accent rating.

Further evidence to support a link between production and perception comes from studies of L1 perception. Newman (2003) investigated whether individuals' perception of speech contrasts was correlated with their production of that contrast. Correlations were examined between acoustic measures taken from subjects' perceptual prototypes for a given speech category and their average production of members of that category. Significant correlations were found for VOT among stop consonants, although the voiceless plosives /p/, /t/, and k/ were more highly correlated than the voiced plosives /b/, /d/, and /g/, and for spectral peaks in voiceless fricatives. Perkell et al. (2004) also found a strong correlation between perception and production. Nineteen young adult speakers of American English were recorded repeating the words *cod*, *cud*, *who'd*, and *hood* in a carrier sentence at a normal, clear and fast speaking rate. The same 19 speakers then discriminated and labelled sets of seven synthetic stimuli ranging from *cod* to *cud* and *who'd* to *hood*, based on productions from one male and one female speaker. The results demonstrated that speakers who had discrimination scores above the median produced greater acoustic contrasts than speakers with discrimination scores below the median.

How might the relationship between production and perception be explained? Several theories of speech perception have claimed that there is a strong link between production and perception. Motor Theory (MT; Liberman et al., 1967; Liberman and Mattingly, 1985) claims that listeners perceive speech in terms of articulatory gestures, and proposes that there is a single, shared representation for speech production and perception. The theory claims that listeners recover the neuromotor commands to the articulators, referred to as 'the intended gestures'. These gestures are thought to be mapped onto more-or-less invariant representations that define phonemes in terms of their articulatory gestures.

Direct Realism Theory (DRT; Fowler, 1981, 1986; Best, 1995) also claims that the objects of speech perception are articulatory rather than acoustic events. However, unlike MT, DRT hypothesizes that listeners perceive the actual vocal tract movements or gestures, rather than the neuromotor commands or intended gestures that cause these movements. The theory also differs from MT in that it claims that there is no specialized phonetic module that mediates speech perception. The information in the acoustic signal is thought to be rich enough to specify the gestures that structure the signal. Thus, all a listener need do to decode speech is detect the relevant physical gestures in the signal.

If speech production and perception are linked in the way that MT or DRT proposes, such that listeners perceive speech through the recovery of 'intended gestures' (MT) or vocal tract movements (DRT), then one could imagine that changes in production (e.g., when learning a non-native language or accent) would also have an effect on perception. That is, the acquisition of new articulatory targets to modify one's own accent may directly change how these sounds are perceived. For example, if a native northern English speaker living in the south of England changes his/her production of the vowel in words like *bud* and *cud* to better match how SSBE speakers produce these words, this may cause changes in the way in which he/she perceives this vowel. Thus, as demonstrated in Study 1, after experience of interacting with SSBE speakers, a native northern English speaker may change his/her best exemplar location for this vowel in SSBE accented carrier sentences.

Recent work in brain imaging has given a possible neurological basis to these theories. Rizzolatti and Arbib (1998) demonstrated that motor areas of the brain are active during perception. They showed that neurons in the premotor cortex of monkeys – so-called mirror neurons – responded when the monkey performed a given action. The same neurons responded when the monkey saw a similar action performed by another monkey or human. Fadiga et al. (2000, 2002) presented evidence for a similar neurological process in humans. In a study using transcranial magnetic stimulation, they showed an increase of motor-evoked potentials in listeners' tongue muscles when they heard words that are produced with strong tongue movements (Fadiga et al., 2002).

However, the experimental evidence for such strong links between perception and production is mixed. Bailey and Haggard (1973, 1980) found no correlation between average VOTs produced in voiced and voiceless consonants and listener's perceptual category boundaries for a /g/-/k/ continuum. Ainsworth and Paliwal (1984) also found a similar result in an experiment investigating listeners' production and perception of English glides. They had listeners produce and identify synthetic tokens of the English glides, /w, r, l, j/. F2 and F3 at the onset of the CV transition were measured in both the production and perceptual identification task. Again, no significant correlations between the production and perception values were found, and frequency transformations from Hertz to mel and Hertz to bark scales, which are thought to better represent how the ear processes the speech signal, also failed to make the correlations statistically significant.

The perception of speech stimuli has also been shown to be similar to that of nonspeech stimuli (e.g., Stevens and Klatt, 1974; Pisoni, 1977), further suggesting that there may not be such a strong link between speech production and perception. Furthermore, studies of nonhuman animals who are unable to produce human speech have shown that they are able to process speech sounds similarly to humans (e.g., Kuhl and Miller, 1975, 1978). For example, Kuhl and Miller (1978) trained chinchillas to respond differently to two endpoint stimuli of a synthetic VOT series ranging from /da/ to /ta/. The animals were then tested with stimuli at intermediate values. Their identification performance corresponded closely with that of adult English-speaking listeners. Further generalization tests with labial (/ba/-/pa/) and velar (/ga/-/ka/) VOT stimuli, as well as tests of VOT discriminability (Kuhl 1981), also showed close agreement with the performance of English speakers.

This evidence supports the view that speech sounds are perceived using general auditory mechanisms, and that speech perception is neither equivalent to nor mediated by the perception of gestures (Diehl et al., 2004). Likewise, changes in perception as a result of experience with a non-native language or accent may not be preceded by or accompanied by changes in production, because the two domains are thought to be independent of each other. Thus, a northerner living in the south of England may change his/her best exemplar locations for the vowels in words like *bud* and *cud* as a result of experience of interacting with SSBE speakers, but this may occur independently of any changes in production.

5.3. Hypotheses

This study examines whether subjects change their production after experience of living in a multidialectal environment where they regularly come into contact with speakers of a non-native accent, and whether this affects their perceptual processes, i.e., their vowel categorization decisions and their ability to recognize a non-native accent in noise.

Cross-language research has emphasized the role of early experience in the ability to accurately perceive and produce foreign or foreign-accented speech (see Section 2.3; e.g., Iverson et al., 2003; Kuhl, 2000). It is possible that the ability to adapt to a non-native accent within the same language is similarly affected; subjects at this late stage in their linguistic development (17-18 years old) may be unable to easily change their production to better match SSBE speakers, even after extensive experience with SSBE. Perception may also be similarly affected; subjects may not be able to easily change their best exemplar locations in SSBE accented-sentences to match those of SSBE speakers and may show no difference in spoken word recognition in SSBE.

Equally, it is possible that after a year's experience of living in a multidialectal environment, subjects will change their production and perception to better match SSBE speakers. Study 1 demonstrated that northerners who had had experience of living in a multidialectal environment changed their best exemplar locations for vowels produced in SSBE accented carrier sentences, and that these best exemplar locations matched descriptions of how northerners produce these vowels when interacting with SSBE speakers (see Chapter 2; Section 2.3). These changes also seemed to be affected by sociolinguistic factors, i.e., the desire to fit in with a new community whilst maintaining one's regional identity; subjects in Study 1 showed changes in their best exemplar locations for the vowels in *bud* and *cud*, but not *bath*. Thus, after experience of living in a multidialectal environment, these subjects may change their production and their best exemplar locations for the vowels in *bud* and *cud*, but retain their native pronunciation of words like *bath*. Likewise, these changes may also affect spoken word recognition: Subjects who change their production and best exemplar locations to better match SSBE speakers may be better at identifying SSBE speech.

Lastly, it is possible that subjects may make changes in one domain but not the other. Evidence from studies of L2 acquisition indicates that production of L2 sounds is easier and/or faster to learn than perception of the same sounds (e.g., Caramazza et al., 1973; Tees and Werker, 1984; Flege and Bohn, 1997). For example, Bohn and Flege (1977) found that for native German speakers, experience with English had a greater effect on the production than the perception of a new vowel category. Similarly, after a year's experience of living in a multidialectal environment, subjects may make greater changes to their production to better match SSBE speakers, but there may be little or no change in their best exemplar locations and their ability to recognize SSBE speech in noise.

Equally, subjects may make greater changes to their perception but not their production. Other studies of L2 acquisition have hypothesized that the accurate production of L2 sounds may lag behind their perception. For example, Flege and Hillenbrand (1984) examined the production of the French syllables /tu/ "tous" and /ty/ "tu" by native speakers of American English. They found that experienced learners produced /y/ like native French speakers, but were unable to produce French /u/ accurately. Instead, they produced this vowel with an F2 that corresponded to the F2 they used to produce French /y/, a value similar to their native English /u/ category. Flege and Hillenbrand (1984) hypothesized that this was because they had not learned to distinguish between French /y/ and /u/. Similarly, subjects in this experiment may change their best exemplar locations to better match SSBE speakers and these changes may affect their ability to recognize SSBE speech in noise, but these changes may not affect their production.

5.4. Experimental Design

A longitudinal study investigated if subjects from the north of England, who had no previous experience of living in a multidialectal environment, changed their speech production and perception when attending university. In Britain, it is usual for students to attend university in a different area to the one in which they have been raised. Consequently, students come into contact with speakers of a wide variety of accents. In particular for this study, students from the north of England come into contact with speakers of southern English accents, including SSBE. In order to fit in with their university community, students change their accent. Specifically then, this study aimed to determine if subjects in early adulthood modified their speech production after experience of living in a multidialectal environment, and whether changes in production were linked to any changes in subjects' perceptual processes, i.e., whether subjects chose different best exemplar locations in a non-native accent after experience with that accent, whether these changes corresponded with any changes in production, and whether these changes affected their ability to process speech in noise produced in the non-native accent.

Students were tested before beginning university (Time 1, T1) 3 months later (Time 2, T2) and on completion of their first year (Time 3, T3). At each testing session they completed three experiments. Experiment 1 (Section 6.1) investigated whether subjects changed their speech production as a result of experience of living in a multidialectal environment. The experiment also included a sociolinguistic interview, conducted at T3. This was included in order to investigate more closely whether there was a link between sociolinguistic factors, such as the desire to fit in with a particular community, and changes in production. Experiment 2 (Section 6.2) used the Goodness Optimization Task (see Study 1: Section 4.2.2) to investigate whether subjects changed their vowel categorization processes, and whether these changes were linked to changes in speech production. Experiment 3 (Section 6.3) investigated whether changes in subjects' accents and vowel category representations had an effect on word recognition. This experiment was included as it has been shown that information learned about a novel sound category within the same language at a segmental, prelexical level is transferred to the lexicon, benefiting word recognition (Eisner and McQueen, in press). It is thus plausible that subjects who change their vowel categorization processes to better match SSBE speakers, will also show changes in their word recognition processes such that they are better able to identify SSBE speech.

6. Study 2: The Experiments

6.1. Experiment 1: Measurement of Production

6.1.1. Introduction

Recording a sample of a speaker's native accent is often problematic. The vernacular, in which the minimum attention is given to speech (e.g. Labov, 1972; Milroy, 1987a), provides the most accurate representation of a speaker's accent because of its highly regular character. However, this is difficult to elicit in experimental situations, as speakers tend to monitor their speech to avoid non-standard variants. In order to elicit the vernacular, spontaneous speech tasks, such as the discussion of a controversial topic, are often used. These tasks encourage the speaker to become engrossed in what they are saying rather than how they are saying it. These tasks can be problematic, though; the target speech sounds may not be produced, or a lot of data may need to be recorded before examples of the target speech sounds that are suitable for analysis are produced.

However, it has been found that subjects do use non-standard variants in tasks typically thought to elicit so-called standard variants. Evans (2001) found that speakers of a northern variety of English used their native, northern vowels when reading a word list, a task typically thought to elicit standard variants. I decided to use similar tasks in this study, as this was considered to be the most efficient way of obtaining a sample of subjects' speech that would be suitable for acoustic analysis and would contain all the phonetic variables of interest. Subjects completed two tasks; they recorded a series of words in a carrier sentence and read a phonetically balanced passage. I collected the data and am originally from Ashby and known to the subjects as a "friend-of-a-friend". Such a relationship has been shown to encourage the use of the vernacular, even in a formal experimental situation (Milroy, 1987b).

Subjects were also interviewed about their attitudes to regional accent variation. Previous work in sociophonetics (e.g., Foulkes and Docherty, 1999) and second language acquisition (e.g., Piske et al., 2001) has emphasized the role of sociolinguistic factors, such as the desire to be identified with a particular community and motivation to learn, in adaptation to a non-native accent or language. It was hypothesized that these factors might also affect the degree to which subjects changed their accent when attending university.

6.1.2. Method

6.1.2.1. Subjects

Twenty-seven subjects were tested and paid for their participation. All were native, monolingual English speakers and reported no speech, hearing or language difficulties. Due to recruitment constraints, the sample group was not balanced for sex; 7 male and 20 female subjects were tested.

Subjects were recruited from Ashby. At the time of recruitment, subjects were aged 17-18 years, and were completing their school education at Ashby Grammar School, the local comprehensive school. All subjects had lived in Ashby since age 5 years, and had been educated at local schools. All subjects had parents and immediate family local to the area, minimizing the risk that subjects had had regular contact with speakers of southern English accents. Subjects attended a range of universities (see Table XI).

Two subjects did not complete the experiment because they dropped out of university during their first term. Of the remaining 25 subjects, two subjects were dropped from the experiment because their best exemplar locations for the

Table XI Universities attended by subjects and the percentages of state-school educated students and students from lower socioeconomic groups at each university. Subjects are referred to by a number preceded by 'M' for males and 'F' for females.

Subject	University attended	% state school pupils	% social class IIIM, IV, V [3]
F01	Bournemouth	91	21
F02	Birmingham	74	13
F03	Leeds	71	16
F04	Lancaster	89	22
F05	York	80	15
F06	Bristol	57	11
F07	Birmingham	74	13
F08	Leicester	85	21
F09	Bristol	57	11
F10	Loughborough	83	22
F11	Hull	89	24
F12	Durham	62	13
F13	York	80	15
F14	Loughborough	83	22
F15	Harpur Adams	77	10
F16	London School of Economics	58	13
M01	Oxford	51	9
M02	Oxford	51	9
M03	Sheffield	81	17
M04	Sheffield	81	17
M05	University of Central England, Birmingham	95	37
M06	Harpur Adams	77	10
M07	Durham	62	13

[3] These groupings are based on the National Statistics Socioeconomic Classification system. IIIM refers to skilled manual workers, IV to partly skilled workers and V to unskilled workers.

goodness optimization task at T1 (Experiment 2; Section 6.2) were not reliable. This gave a test sample of 23 subjects, 7 male and 16 female subjects. As in Study 1, subjects' results were considered to be unreliable if their best exemplar locations for vowels that are produced in the same way in southern and northern English (*bead*, *bed*, *bird* and *bad*) accents differed by more than 2 ERB.

6.1.2.2. Stimuli and Apparatus

The stimuli consisted of eleven test words in the carrier sentence *I'm asking you to say the word [ ] please*, and a phonetically-balanced passage, 'Arthur the Rat' (Appendix D). The test words were the same as those for which subjects found best exemplars in the Goodness Optimization Task (Experiment 2; Section 6.2); *bad, bard, bawd, bed, bird, booed, bud, bead, cud, could,* and *bath*. This allowed potential changes in subjects' best exemplar locations to be compared with changes in their production.

All recordings were made in a quiet room in the researcher's or subject's home using a Sony DAT (Digital Audio Tape) recorder and Sony microphone.

6.1.2.3. Procedure

Recording. The test words were printed in the carrier sentence on separate cards. Subjects were instructed to read each sentence aloud at a normal pace. Subjects recorded two repetitions of each target word in a randomized order. 'Arthur the Rat' was printed on a piece of card. Subjects were instructed to read the passage as if they were reading it to the researcher. It was thought that this would encourage subjects to speak in the way that they would when interacting with friends at university.

Sociolinguistic interview. At the end of the final testing session, subjects were interviewed about their attitudes to accent. The procedure was based on that developed by Sangster (2002) in her study of accent adaptation in students from Liverpool attending Oxford University. Subjects were asked to read through a passage entitled "A study of young people from Tyneside" (Watt, 1998; see Appendix E), which discusses the dilemmas of having a Tyneside or "Geordie" identity that young people might face. Subjects were asked to comment on and discuss the ideas put forward in the passage, and how these related to their own experiences. This passage was used because, as Sangster (2002) explains, it discusses specific and conflicting attitudes that young people might have to their background without explicitly mentioning accent. The passage encourages subjects to think about how they regard their 'northern' background, and how aspects of this might conflict with their desire to fit into a new, multidialectal community.

Accent Ratings. Five phonetically trained listeners rated samples of subjects' accents. The ratings were based on a short sample of subjects' recordings of 'Arthur the Rat' at T1 (i.e., before subjects' began university) and T3 (i.e., after they had completed one year at university);

Just then the old captain saw Arthur. "Stop," he ordered the others coarsely. "You are coming, of course?" "I'm not certain," said Arthur, undaunted. "The roof may not come down yet."

This passage was selected because it was thought to be revealing of accent. In particular, it contains examples of the vowels that were expected to change as a result of experience with SSBE speakers; the vowels in words like *just, coming, come.*

There was one testing session for each rater lasting approximately 30 minutes. The session was self-paced and subjects were told that they could listen to each sample as many times as they wished. Subjects played each stimulus by clicking with a computer mouse on the sound file. Subjects gave their ratings by indicating on a response sheet, on a scale of one to ten, how northern or southern they thought the speaker sounded to them. A response of '10' meant that the speaker sounded 'very northern' and a response of '1' meant that the speaker sounded 'very southern'. Subjects were not aware of the design of the study: they were not told that they would hear the same speaker more than once, and were unaware that speakers were expected to have changed their accent. Subjects rated 44 samples. Due to problems with the recordings, there were no suitable stimuli available for two speakers at T1 (M05 and F09). The order of presentation was randomized across participants.

Acoustic analysis. The acoustic analysis was based on measurements of the test words produced in the carrier sentence *I'm asking you to say the word [ ] please*. F1, F2 and duration were measured for each test word, giving two sets of measurements for each target word. F1, F2 and duration were averaged for each word and these measurements were used in all subsequent analyses.

Acoustic measurements were made in Praat (Boersma and Weenink, 2004). Stimuli were located manually, and then F1 and F2 were measured using the formant tracker. Formant frequencies were measured from the mid-point of the steady-state portion of the vowel. The steady-state portion of the vowel was defined as the point at which the formants were parallel to the time axis of the spectrogram, and thus not changing (Clark &Yallop, 1996). Where the formant tracker was unable to measure the formant accurately (e.g., because of poor resolution), the formant measurement was taken manually. Duration was measured manually from a spectrogram. All duration measurements were taken from the beginning of the F2 transitions to the end of the F2 transitions.

So that the data from male and female speakers could be compared, a version of Nearey's (1978) individual log mean procedure was used to normalize the production data. In this procedure, each log-transformed formant frequency is expressed as a distance to a reference point, the log mean. This procedure was chosen because it has been shown to be one of the most effective methods of reducing the amount of anatomical and physiological variation, while retaining sociolinguistic and phonemic variation (Adank, 2003; Adank et al., 2004). The average formant frequency measurements for each experimental word were normalized following the equation:

$$F_{ijk\ norm} = G_{ijk} - G \qquad (4)$$

where i is the formant, j is the vowel being transformed, k is the speaker, G is the log-transformed frequency of formant i, and G is the log mean for a speaker k. In this study the log mean was the mean F1 and F2 of the vowels that are produced similarly in northern and southern English accents in the words in sentences for that talker (i.e., *bad*, *bard*, *bawd*, *bead*, *bed*, *bird*, and *booed*), averaged over T1, T2, and T3. An average was used to minimize the risk that random variation in the production data might affect the normalization procedure. The vowels in *bud*, *cud*, *could*, and *bath* were omitted as these vowels were expected to change, and it was thought that including these vowels might also reduce the accuracy of the normalization procedure.

6.1.3. Results

6.1.3.1. Perceived Accent Rating

Before the ratings were examined for changes in subjects' accent, it was necessary to establish that the ratings were reliable (i.e., that raters were using the scale in the same way). A Pearson correlation between all pairs of raters demonstrated that subjects' accent ratings were in the range of 0.458 to 0.677, confirming that the ratings had a significant level of agreement. Consequently, the ratings were averaged across raters and these values were used in all subsequent analyses.

As displayed in Table XII, 21 out of 23 subjects were rated as sounding more southern at T3 than T1 (i.e., they were given a lower rating). The potential differences between the average ratings at T1 and T3 were tested in a repeated measures ANOVA, with time (T1 or T3) coded as a within-subject variable. There was a main effect of time, $F(1,20) = 24.84, p < 0.001$, confirming that the accents of the subjects had changed to become more southern. The effect of time on accent ratings can be clearly seen in Fig. 7. The average rating at T3 is lower than at T1, demonstrating that subjects were judged to sound more southern at T3.

However, the differences between the accent ratings for each subject at T1 and T3 were small (Fig. 8): 18 of the 23 subjects tested changed their accent by less than 2 points on the rating scale, and the remaining subjects changed their accent by less than 3 points on the rating scale (Table XII).

In addition to accent change, there were large individual differences in the overall accents of the talkers; some subjects were judged to have a particularly northern accent at both T1 and T3, but others were judged to have a particularly southern accent (Fig. 8). Although 20 out of 23 subjects had an accent rating at T1 and T3 of between 7 and 5, one subject was given a rating of above 8 at T1 and T3 and two subjects were given a rating of less than 5 at T1 and T3. These differences can be seen clearly when the average rating over T1 and T3 for each subject – the overall accent rating – is compared (Table XII). For example, M07 had an overall accent rating of 3.8, but M06 had an overall accent rating of 8.9 on the rating scale.

Table XII Accent ratings for all subjects (N=23).The table shows the five listeners' ratings (R1-R5) for each subject at T1 and T3, the average rating at T1 and T3, the overall accent rating (the average of all the ratings at T1 and T3 for that subject) and the change in accent rating (the average accent rating at T3 minus the average accent rating at T1 for each subject).

Subject	T1					Average (T1)	T3					Average (T3)	Overall Accent Rating	Change in Accent Rating
	R1	R2	R3	R4	R5		R1	R2	R3	R4	R5			
F01	7	7	8	7	8	**7.4**	7	6	9	6	8	**7.2**	**7.3**	**-0.2**
F02	8	6	9	8	6	**7.4**	8	6	9	5	8	**7.2**	**7.3**	**-0.2**
F03	8	6	8	8	8	**7.6**	6	6	7	6	8	**6.6**	**7.1**	**-1**
F04	6	5	8	6	6	**6.2**	6	3	6	3	2	**4**	**5.1**	**-2.2**
F05	8	5	7	6	6	**6.4**	7	4	6	5	4	**5.2**	**5.8**	**-1.2**
F06	7	8	9	8	10	**8.4**	7	6	8	7	9	**7.4**	**7.9**	**-1**
F07	7	6	7	6	6	**6.4**	7	5	8	6	6	**6.4**	**6.4**	**0**
F08	7	5	6.5	5	8	**6.3**	6	4	6	5	6	**5.4**	**5.85**	**-0.9**
F09	-	-	-	-	-	**-**	4	3	4	1	1	**2.6**	**2.6**	**-**
F10	8	6	8	7	9	**7.6**	7	7	7	5	8	**6.8**	**7.2**	**-0.8**
F11	6	7	7	4	6	**6**	7	6	8	6	6	**6.6**	**6.3**	**0.6**
F12	8	5	5.5	6	6	**6.1**	5	4	8	6	5	**5.6**	**5.85**	**-0.5**
F13	7	6	7	6	7	**6.6**	6	4	6	5	7	**5.6**	**6.1**	**-1**
F14	6	5	8	5	6	**6**	6	5	7	5	7	**6**	**6**	**0**
F15	7	4	7	9	9	**7.2**	6	3	7	5	7	**5.6**	**6.4**	**-1.6**
F16	6	3	7	5	7	**5.6**	6	2	6	6	4	**4.8**	**5.2**	**-0.8**
M01	7	7	8	8	9	**7.8**	6	3	8	4	7	**5.6**	**6.7**	**-2.2**
M02	8	7	10	7	9	**8.2**	6	5	7	4	7	**5.8**	**7**	**-2.4**
M03	8	9	9	5	8	**7.8**	6	4	7	6	8	**6.2**	**7**	**-1.6**
M04	6	5	5	4	7	**5.4**	6	4	9	7	3	**5.8**	**5.6**	**0.4**
M05	-	-	-	-	-	**-**	7	9	7	7	10	**8**	**8**	**-**
M06	9	10	10	8	10	**9.4**	8	7	10	7	10	**8.4**	**8.9**	**-1**
M07	3	6	7	3	3	**4.4**	2	3	6	3	2	**3.2**	**3.8**	**-1.2**

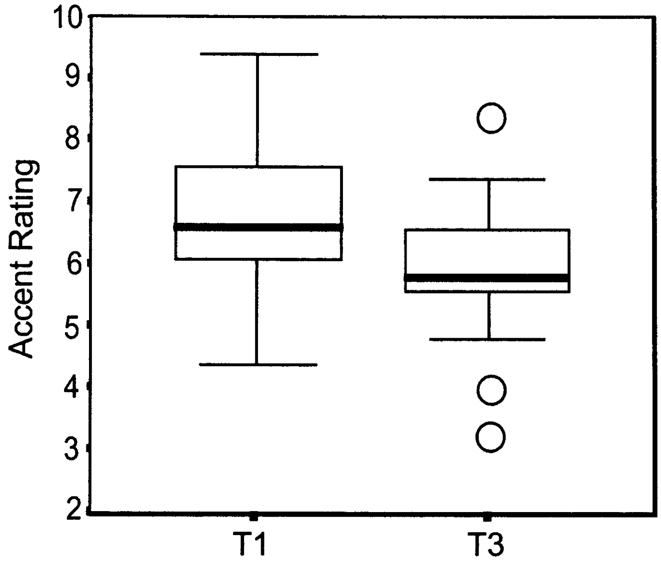

Figure 7 Boxplot of the accent ratings for subjects before beginning university (T1) and on completion of their first year at university (T3).

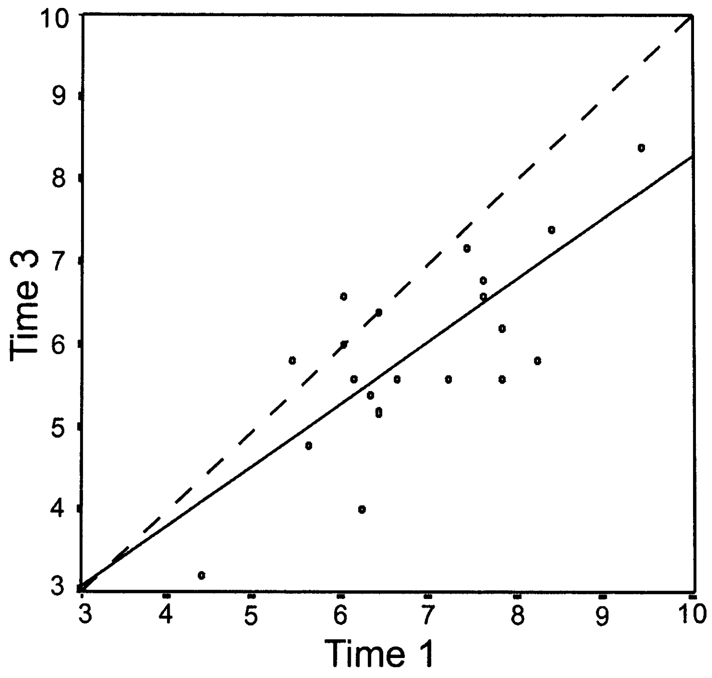

Figure 8 Scatterplot of the average accent rating for subjects at T1 and T3. The dotted line is the line of no change. The solid line is the line of best fit. Although most subjects were judged to have changed their accent (i.e., all those below the dotted line), the changes were often small. There was also some variability in the data; although most subjects were given an accent rating between 6 and 8, some subjects were given a particularly low (i.e., southern) rating, and others were given a particularly high (i.e., northern rating) at both T1 and T3.

6.1.3.2. *Acoustic analysis*

As displayed in Fig. 9, subjects changed the way in which they produced the vowels in *bud*, *cud*, and *could*. At T1 subjects produced these words with a high-back vowel, as is typical for speakers of northern English accents (see Chapter 2). At T2 some, but not all, subjects had begun to centralize these vowels. At T3, almost all subjects produced these words with a more central vowel so that they were acoustically similar to the vowel in *bird*. The shift appeared to occur predominantly in the F2 dimension, although there were also changes in the F1 dimension; some subjects produced *bud* and *cud* with a higher F1 at T3 (i.e. lower in the vowel space). There was little change in F1 and F2 for all other vowels. As displayed in Table XIII, there also appeared to be little change for duration.

A Pearson correlation examined whether the observed changes in F1 and F2 for *bud*, *cud*, and *could* predicted the change in accent ratings (i.e., the difference in accent ratings at T1 and T3). *Bath* was also included in the analysis because even though subjects did not appear to change their production of this word, it is produced differently in southern English accents. The correlation demonstrated that the changes in F1 and F2 for *bud* but not *cud* were significantly correlated with the change in accent rating (Table XIV), confirming that subjects who changed their accent to sound more southern changed their production of the vowel in *bud* so that it was closer to how southerners produce this vowel. There were no significant correlations for any other vowels and the change in accent rating.

The accent ratings analysis demonstrated that there were individual differences in production; some subjects were judged to have a more southern accent at both T1 and T3 than others. A Pearson correlation was used to investigate whether the individual differences in subjects' overall accent ratings predicted individual

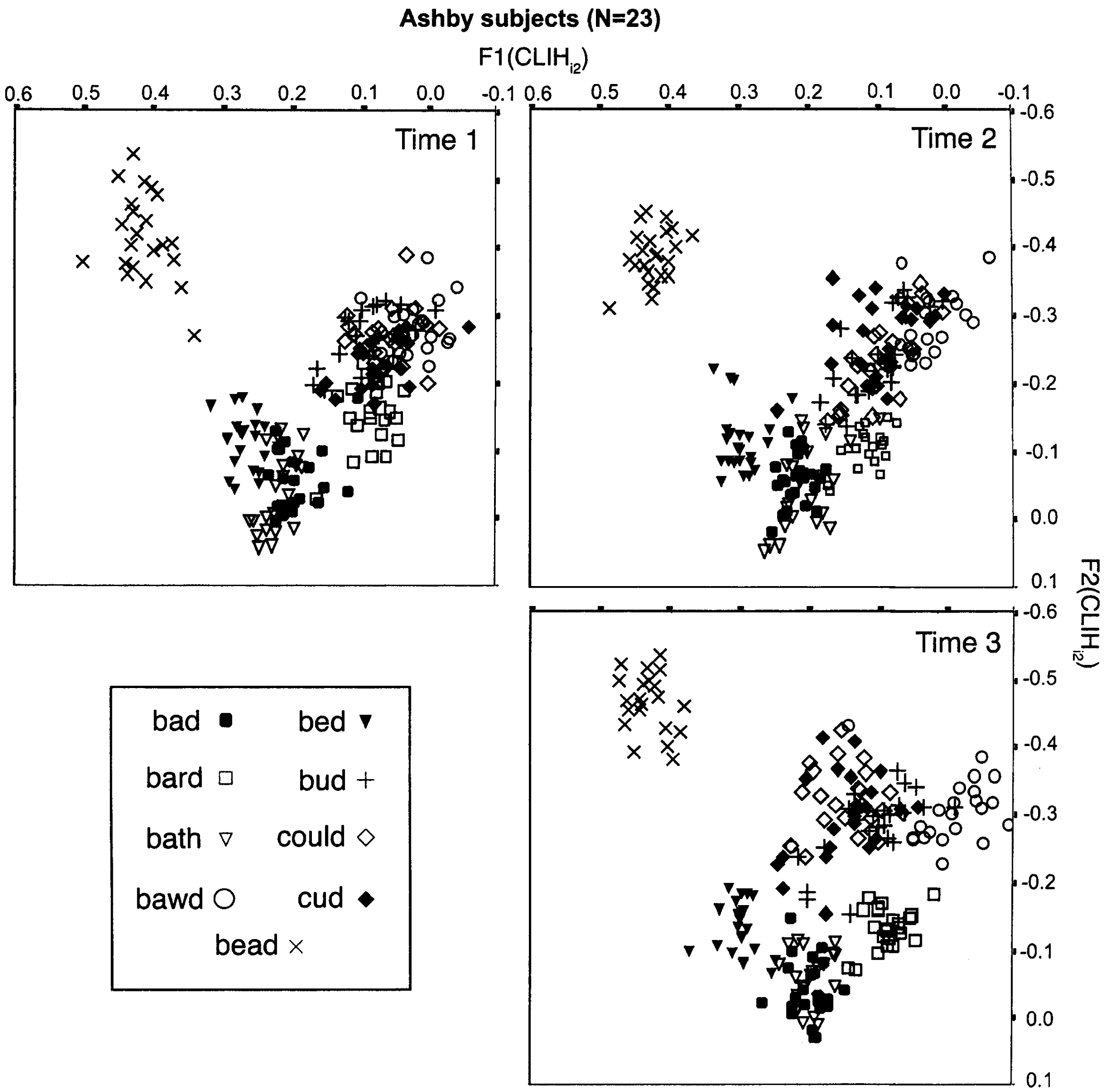

Figure 9 Normalized F1 and F2 formant frequencies (CLIH$_{i2}$) of subjects' productions of target words in the carrier sentence *I'm asking you to say the word [ ] please*, at T1, T2 and T3. Subjects appeared to change their production of *bud*, *cud*, and *could* as a result of experience with SSBE speakers. This change appeared to occur predominantly in the F2 dimension; subjects produced these words with a more fronted vowel (i.e., higher F2) at Time 3. Some subjects also changed F1; they produced these words with a lower vowel (i.e., higher F1).

Table XIII Average durations (ms) of vowels in the target words *bud*, *cud*, *could*, and *bath* for each subject at T1, T2, and T3.

Subject	bud			cud			could			bath		
	T1	T2	T3	T1	T2	T3	T1	T2	T3	T1	T2	T3
F01	157	148	160	135	120	131	158	163	162	150	153	152
F02	140	133	135	110	103	105	110	115	112	100	101	93
F03	143	138	141	140	147	144	138	127	124	130	132	126
F04	115	106	109	99	92	93	102	97	99	80	101	75
F05	153	152	157	108	115	112	111	106	108	108	114	109
F06	119	125	123	100	91	92	115	120	119	89	103	100
F07	154	148	150	147	149	147	141	140	143	158	153	157
F08	158	-	161	159	153	158	168	159	164	100	103	104
F09	-	-	-	-	-	-	-	-	-	-	-	-
F10	106	105	107	110	95	90	103	117	105	104	107	107
F11	141	137	139	121	120	122	147	149	144	168	162	164
F12	112	121	117	111	118	114	121	115	119	100	98	106
F13	159	163	161	149	141	138	133	135	133	117	115	118
F14	117	122	121	110	103	107	103	98	100	119	123	120
F15	141	142	137	75	79	76	101	94	95	125	119	122
F16	135	129	131	129	134	133	118	115	116	157	152	153
M01	160	159	177	136	133	135	158	149	155	179	171	176
M02	124	126	126	119	125	122	111	115	110	129	131	127
M03	129	134	130	116	114	112	121	124	123	121	120	123
M04	124	128	126	111	119	118	132	129	130	104	109	110
M05	-	-	-	-	-	-	-	-	-	-	-	-
M06	124	126	123	102	104	102	117	121	116	108	103	110
M07	138	131	139	117	115	114	106	108	108	103	101	104

Table XIV Correlation matrix of the change in accent rating and change in F1, F2 and duration for *bud*, *cud*, *could* and *bath* for all subjects (N=23). The changes in F1 and F2 for *bud* were significantly correlated with the change in accent rating. There were no significant correlations for any other vowels.

	Change in Accent Rating	
Bud: F1	Pearson Correlation	-0.566
	Sig. (2-tailed)	**0.007****
Bud: F2	Pearson Correlation	-0.464
	Sig. (2-tailed)	**0.034***
Bud: Duration	Pearson Correlation	0.145
	Sig. (2-tailed)	0.531
Cud: F1	Pearson Correlation	-0.291
	Sig. (2-tailed)	0.201
Cud: F2	Pearson Correlation	-0.292
	Sig. (2-tailed)	0.198
Cud: Duration	Pearson Correlation	0.199
	Sig. (2-tailed)	0.388
Could: F1	Pearson Correlation	-0.146
	Sig. (2-tailed)	0.529
Could: F2	Pearson Correlation	-0.193
	Sig. (2-tailed)	0.401
Could: Duration	Pearson Correlation	0.097
	Sig. (2-tailed)	0.676
Bath: F1	Pearson Correlation	-0.313
	Sig. (2-tailed)	0.167
Bath: F2	Pearson Correlation	-0.136
	Sig. (2-tailed)	0.557
Bath: Duration	Pearson Correlation	-0.004
	Sig. (2-tailed)	0.985

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

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

differences in the production of *bud*, *cud*, *could*, and *bath* at T1, T2, and T3. F1, F2, and duration for *bud*, *cud*, *could*, and *bath* were averaged over T1, T2, and T3 for each subject and entered into the analysis with the overall accent rating for each subject. The correlation demonstrated that the overall accent rating was significantly correlated with subjects' average F1 and F2 for *bud* and *cud* (Table XV), confirming that subjects who were judged to have a more southern accent produced these words with a more southern vowel. There were no significant correlations for any other vowels and the overall accent rating.

In summary, subjects were judged to have changed their accent as a result of experience of living in a multidialectal environment, although the change was often small. There were also individual differences in production: some speakers were judged to have a more southern accent than others. Pearson correlations demonstrated that the change in accent rating and the overall accent rating were related to changes and differences in the way in which subjects produced the vowels in *bud* and *cud*. Subjects who changed their accent to sound more southern changed their production of the vowel in *bud* and *cud* so that it was closer to how southerners produce this vowel, and subjects who sounded more southern overall produced these words with a more southern vowel at T1, T2, and T3.

Table XV Correlation matrix of the overall accent rating and the overall average F1, F2 and duration for *bud*, *cud*, *could* and *bath* for all subjects (N=23). F1 and F2 for *bud* and *cud* were significantly correlated with the overall accent rating. There were no significant correlations for any other vowels.

	Overall Accent Rating	
Bud: F1	Pearson Correlation	-0.666
	Sig. (2-tailed)	**0.001****
Bud: F2	Pearson Correlation	-0.568
	Sig. (2-tailed)	**0.005****
Bud: Duration	Pearson Correlation	0.176
	Sig. (2-tailed)	0.446
Cud: F1	Pearson Correlation	-0.638
	Sig. (2-tailed)	**0.001****
Cud: F2	Pearson Correlation	-0.508
	Sig. (2-tailed)	**0.013***
Cud: Duration	Pearson Correlation	0.269
	Sig. (2-tailed)	0.238
Could: F1	Pearson Correlation	-0.004
	Sig. (2-tailed)	0.987
Could: F2	Pearson Correlation	-0.346
	Sig. (2-tailed)	0.106
Could: Duration	Pearson Correlation	0.157
	Sig. (2-tailed)	0.498
Bath: F1	Pearson Correlation	0.104
	Sig. (2-tailed)	0.637
Bath: F2	Pearson Correlation	-0.067
	Sig. (2-tailed)	0.760
Bath: Duration	Pearson Correlation	0.105
	Sig. (2-tailed)	0.649

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

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

6.1.3.3. Sociolinguistic interview

This section reviews the data collected in the sociolinguistic interview. The section reports subjects' attitudes to their own accent and regional accent variation in light of their experiences at university, and discusses how differences in attitude might be able to account for differences in the amount of accent change and individual differences in accent found in the acoustic analysis of the production data.

Subjects attended a range of universities that differ in the type of students they attract (Table XI, Section 6.1.2.1). For example, universities such as Oxford tend to attract more students from the upper social classes who have been privately educated than universities such as the University of Central England that tend to attract students from lower socioeconomic backgrounds who have been educated at state schools. SSBE is typically associated with upper-middle and middle class speakers and, due to its connotations of prestige and education, is particularly common amongst educated speakers[4]. Subjects who attended universities like Oxford

---

[4] It might be argued that RP (Received Pronunciation) would be a more accurate description of this accent. Although some speakers at universities like Oxford might be described as speakers of RP, or near-RP (e.g., Wells, 1982), this group likely makes up only a small percentage of Oxford's student community, with most being drawn from the upper-middle and middle classes. Members of these classes are not traditionally regarded as RP speakers; although these speakers may be described as having an "educated" or standard accent, they often acquire this accent later in life and use standard forms alongside regionalisms (e.g., Wells, 1982). Attitudes to RP have also changed (see Fabricius, 2002 for a review); it has recently been described as snobbish (Newbrook, 1999) and is regarded with hostility by some speakers of

are likely to have come into contact with more speakers of SSBE than subjects at universities like the University of Central England. It is possible that these differences affected the amount by which subjects changed their accent. One could imagine that subjects who attended Oxford, changed their accent more than those who attended universities where there were fewer SSBE speakers.

Of the three subjects who changed their accent the most – F04, M01 and M02 - M01 and M02 attended Oxford University. Both subjects reported that their friendship groups were made up predominantly of SSBE speakers:

"some friends are from Yorkshire, but mostly from London". (M01)

"I'm around lots of southern people...my friends they're mostly from the south". (M02)

In contrast, the subjects who did not change their accent or who changed their accent by less than 1 point on the rating scale and had a high overall accent rating (F01, F02, F07, F08, F10, F11 and F14) attended universities with a higher proportion of state-school educated students. These subjects reported that their friendship groups were made up predominantly of speakers of non-standard accents rather than SSBE and, in the case of F11, had a stronger northern accent than themselves:

"they all think I sound well posh but that's strange cos I don't have a posh accent. I'm from the north". (F11)

---

non-standard varieties (Stuart-Smith, 1999). Thus, in my opinion it is more accurate to describe the dominant accent of the university community in Oxford and the accent to which these subjects might aspire as SSBE rather than RP or near-RP.

Some of these subjects also attended university locally. F02, F07, F08, F10, and F14 all attended universities in the Midlands and either lived at home with their parents or returned to Ashby frequently (i.e., at least once a week). It is possible that M01 M02 and F04 changed their accent more because they came into contact with a greater number of SSBE speakers, and that F01, F02, F07, F08, F10, F11, and F14 did not change their accent to the same extent because they had not had as much experience with SSBE speakers. Indeed, subjects who changed their accent emphasized that they thought that the fact that they were in frequent contact with southerners had influenced their accent:

"It's just cos I'm in contact with people who speak properly more, you know what I mean, with an RP or whatever, you know what I mean, not quite that extreme, but like, but they pronounce it more like that". (M01)

"Sometimes I find myself saying stuff like 'laugh' ([lɑ:f]) and Cat (her friend) goes ha ha ha you just said 'laugh' ([lɑ:f]) and I'm like that's from being around you". (F04)

Motivation to fit into a particular community also appeared to play a role in explaining why some subjects changed their accent more than others. M01, F04, F15, and M07 changed their accent by more than 1 point on the rating scale, and all reported that their friendship groups at university were predominantly made up of SSBE speakers. These subjects all reported that they thought their friendship groups had had an influence on their accent:

"Before I went away I didn't think I had an accent as such...but when I went for my interview people pointed out words that I said and sort of laughed at them, things like 'butter', they say 'butter' [ˈbətə] I say 'butter' [ˈbʊtə], so like I didn't feel ashamed but I'm slightly more well-spoken now. I think that's because of their southern influence [laughs]". (M02)

"It (her accent) changed like the first couple of weeks I was there. It got posher. I think that's cos all the people I met have like southern accents". (F14)

It is thus plausible that subjects changed their accent to better fit in with their friendship groups and because they felt that it was important to identify themselves with this new community (cf. Foulkes and Docherty, 1999).

However, the majority of subjects did not change all aspects of their native accent. Instead, there appeared to be a tension between subjects' desire to sound more cosmopolitan and to fit in with a multidialectal community, and their wish to continue to signal loyalty to their native community (see also Foulkes and Docherty, 1999). F04 was the only subject to report using the southern variant [ɑː] in words like *laugh* when interacting with her friends. All other subjects reported changing the way in which they produced words like *butter*, but not *bath*:

"I think it, butter, has slightly changed. Single vowel words like *'duck'* [dʊk] I think have slightly moved away from the 'duck' [dʊk] to 'duck' [dək]. I'm never gonna say 'grass' [gɹɑːs] and 'bath' [bɑːθ], just, I don't know why, it just doesn't sound right. (M02)

"I'll never say 'bath' [bɑːθ]. I think I'll always say 'bath' [bæθ] and [gɹæs]. (...)Well I was talking to somebody the other day and I said 'oh but' [bət]...um so things like that depending on who I'm talking to I'll often say 'but' [bət]. (F04)

Subjects also agreed with the view put forward in the Watt passage (Appendix E); they felt that although northerners wanted to change aspects of their accent to fit in with what they saw to be a more cosmopolitan community, it was important that they retained aspects of their accent that identified them with their native accent community. For example, M01 reported that he felt it was important to fit in and be identified with his southern friends at university, but that he did not want to be accused of "having lost sight of where he had come from". It is thus plausible that this is why subjects changed their production of some words but retained their native productions in others. Subjects may have retained their native vowel in words like *bath* to continue to signal belonging to their native accent community, but have changed the vowel in words like *bud* and *cud* in order to identify themselves with their new, university community.

It is less clear from the interview data why some subjects were given a low overall accent rating (i.e., were judged to sound more southern) and others a high overall accent rating (i.e., were judged to sound more northern). All subjects had been born and raised in Ashby and had attended local schools throughout their childhood. They were also all from very similar socioeconomic and social backgrounds. It is possible that subjects who had a low overall accent rating had changed their accent before beginning university. For example, they may have felt that a SSBE accent was more prestigious and would be advantageous in their future career. However, only one subject, F16, reported this view. She felt that she had had a strong northern accent but that she had begun to change her accent to sound more southern as a teenager. She reported that this was because of her career aspirations. She wanted to be a barrister and felt that having a strong northern accent would mean that she would not be taken seriously and would be regarded as uneducated:

"My Dad told me that if I want to become a barrister you can't talk like you do. If I want to go into the job I want to I have to make a conscious effort to change my accent. Just to sound [.] people associate people from the south with better education, more money, better culture and that's important for me. (F16)

Another possibility is that parental background affected subjects' accent development in these subjects. Although subjects' parents were originally from the local area, it is possible that they had moved away for university or employment before returning to Ashby. These parents may have been exposed to southern English speakers, and have changed their accent as a result. As such, their children may not have acquired or felt that it was important to acquire, a strong local accent.

6.2. Experiment 2: An investigation of best exemplar locations

6.2.1. Introduction

Study 1 (Chapter 4) suggested that there was a link between production and perception, such that in order to adapt perceptually to a non-native accent a listener needed to have experience of changing his/her production when interacting with a native speaker of that accent (see Chapter 4, Section 4.4). This experiment investigated whether subjects changed their vowel categorization decisions after experience of living in a multidialectal environment, and whether these changes were linked to the changes they made in production (i.e., whether subjects who changed their accent to a greater extent changed their vowel categorization decisions more, and whether subjects who had a more southern accent overall, chose more southern vowels).

6.2.2. Method

6.2.2.1. Subjects

The subjects were the same as those tested in Experiment 1.

6.2.2.2. Stimuli and Apparatus

The stimuli were the same as those used in Study 1, Experiment 2 (Chapter 4, Section 4.3). Stimuli were played at a sampling rate of 11 kHz using a computer sound card. Subjects listened over headphones (Sennheiser HD 414) in a quiet room in the subject's or researcher's home.

6.2.2.3. *Procedure*

Subjects found best exemplars for eleven test words: *bad, bard, bawd, bed, bird, booed, bud, bead, cud, could,* and *bath*. Subjects found best exemplars for *bud*, *cud*, and *bath*, because these vowels are produced differently in SSBE and Sheffield English. *Could* was included as it was thought that potential modifications to *bud* and *cud*, might affect subjects' best exemplar location for this word (see Chapter 4, Section 4.2.1.2). *Bard* was included so that any modifications to the *bath* category could be compared with this vowel: Southerners produce both *bath* and *bard* with the same vowel, [ɑ], but northerners produce *bath* using [a], even though they produce words like *bard* using the same vowel as southerners. Subjects also found best exemplars for *booed* and *bawd* to show the height and width of the vowel space respectively. To give a measure of reliability, subjects found best exemplars for *bead*, *bed*, *bird* and *bad*. These vowels are produced similarly in northern and southern English accents, and so subjects should choose the same best exemplar locations in both SSBE and Sheffield English carrier sentences.

The procedure was the same as Study 1, Experiment 2 (Chapter 4, Section 4.2.1.3), in all other respects.

6.2.3. Results

6.2.3.1. *Bud and Cud*

A repeated measures MANOVA (Multivariate Analysis of Variance) analysis tested whether subjects changed their vowel categorization decisions for *bud* and *cud*, and whether these changes were linked to changes in production (i.e., differences in the change in accent rating) or individual differences in production (i.e., differences in the overall accent rating). Word (*bud* and *cud*), time (T1, T2 and T3) and sentence context (SSBE or Sheffield English) were coded as within-subject variables and change in accent rating and overall accent rating were coded as covariates.

For F2, there were no significant main effects or interactions, $p > 0.05$, demonstrating that the shifts in the F2 dimension shown in Figs.10-14 were unreliable. This confirms that subjects were not changing their vowel categorizations in this dimension and that there were no reliable differences between subjects in terms of F2. There were also no significant main effects or interactions for duration, $p > 0.05$. This can be seen in Table XVI: There are few changes in duration over time, and subjects chose similar durations in SSBE and Sheffield English carrier sentences.

For F1, there was a main effect of sentence context, $F(1,17) = 6.16, p < 0.05$, demonstrating that all subjects chose different best exemplar locations for *bud* and *cud* in SSBE and Sheffield English carrier sentences. As displayed in Fig. 10, all subjects chose a more central best exemplar location (i.e., higher F1) in SSBE carrier sentences at T1, T2, and T3. There was also a significant interaction of sentence context and time, $F(1,17) = 8.08, p < 0.05$, indicating that the amount of normalization for these vowels in SSBE and Sheffield English carrier sentences changed over time. This can be seen in Fig. 10; the amount of normalization for *bud* and *cud* decreases over time so that it is smaller at T3 than at T1. A paired samples t-test confirmed that the amount of normalization was greater at T1 than T3, t=2.911, $p < 0.05$. There was no difference in the amount of normalization at T1 and T2, t=0.716, $p > 0.05$, or T2 and T3, t=0.867, $p > 0.05$.

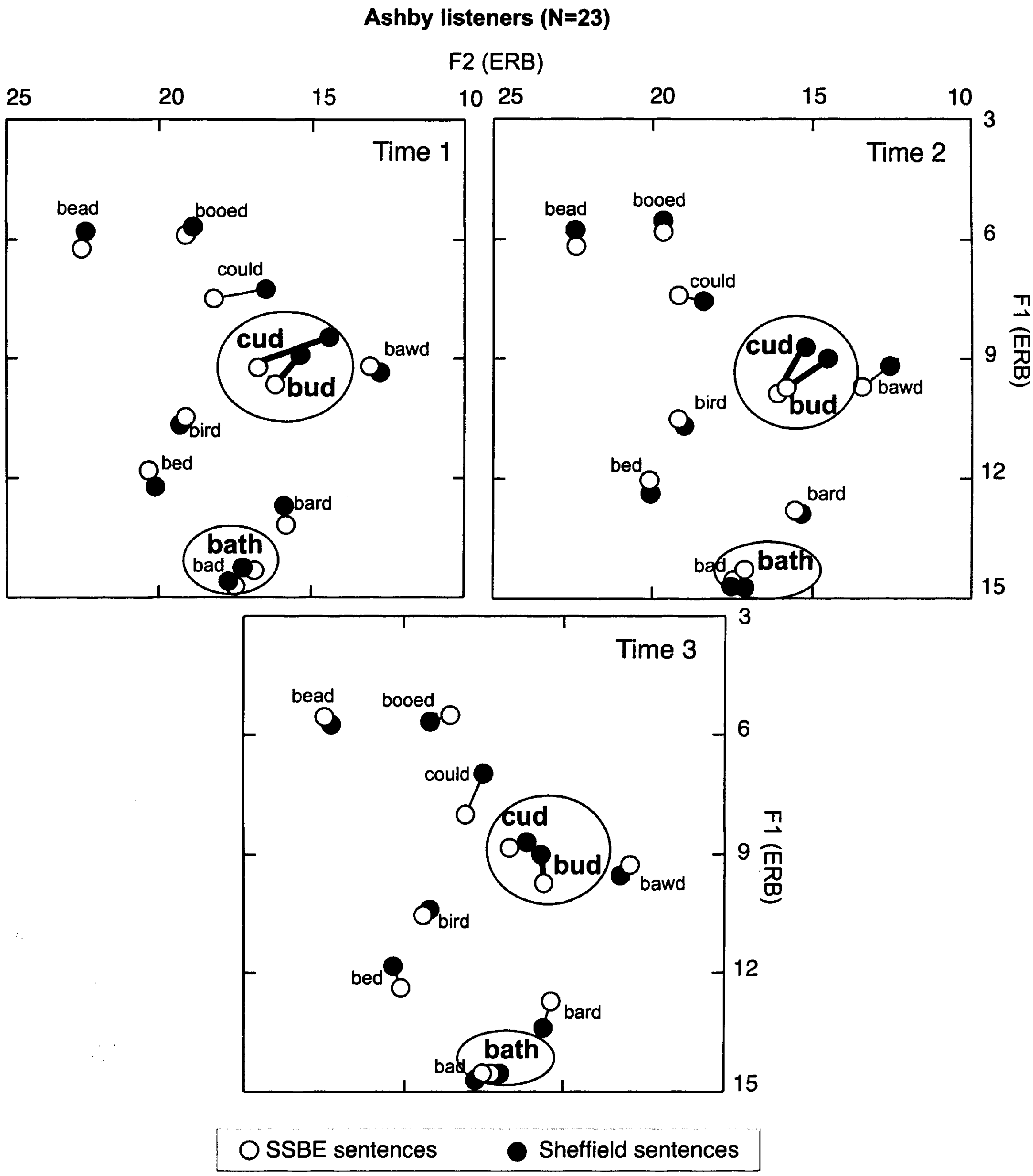

Figure 10 Average F1 and F2 formant frequencies (ERB) of best exemplars for all subjects in SSBE and Sheffield English carrier sentences at T1, T2, and T3. All subjects appeared to choose lower vowels (i.e., higher F1) for *bud* and *cud* in SSBE carrier sentences at T1 and T2, but at T3 they chose lower vowels in both carrier sentences. There were no changes for any other vowels.

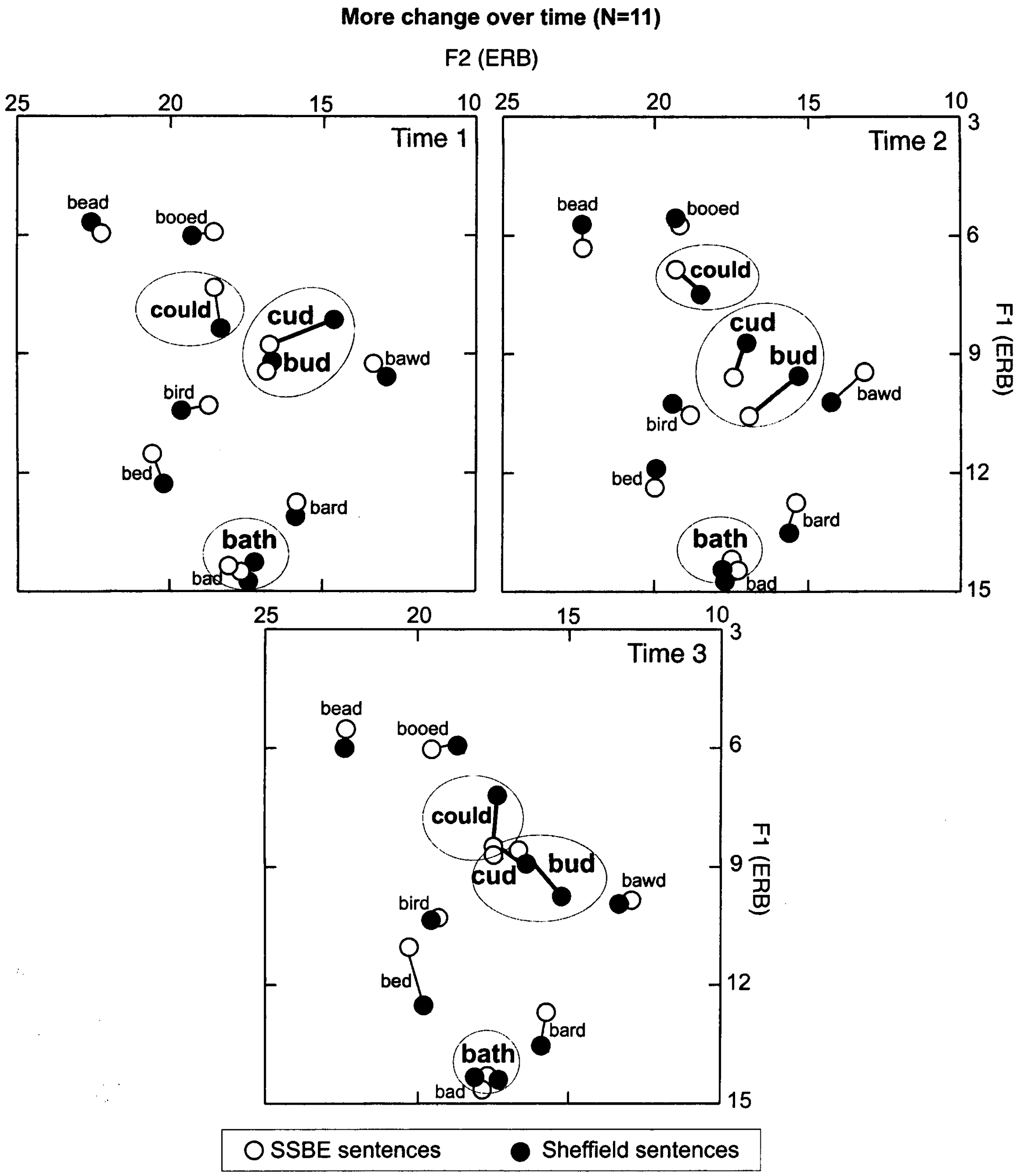

Figure 11 Average F1 and F2 formant frequencies (ERB) for best exemplars in SSBE and Sheffield English carrier sentences for subjects who were judged to have changed their accent more. These subjects changed their best exemplar location for *could* after experience of living in a multidialectal environment; they chose a lower vowel (i.e., higher F1) for *could* at T3. There were no other reliable changes in subjects' best exemplar locations.

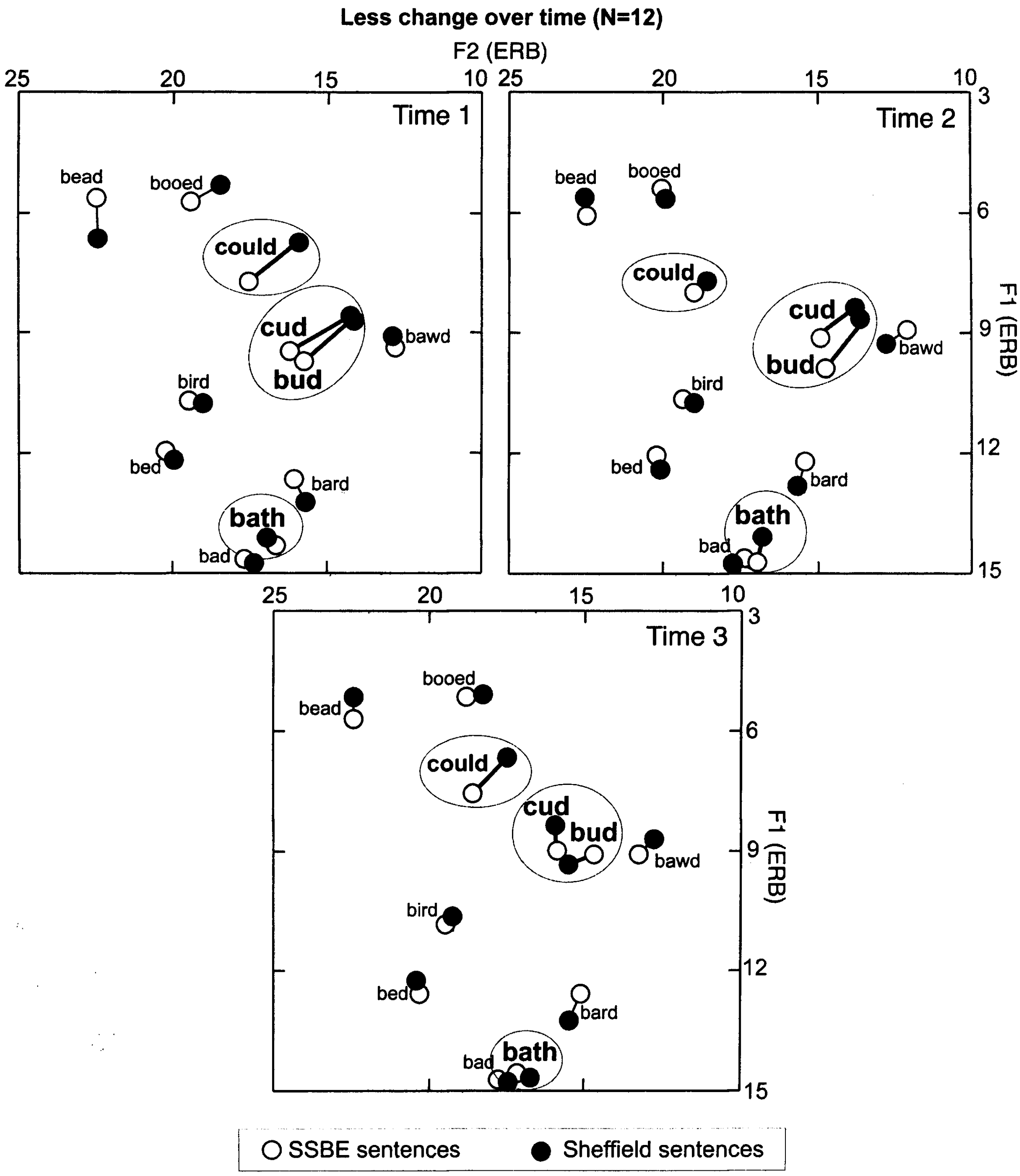

Figure 12 Average F1 and F2 formant frequencies (ERB) for best exemplars in SSBE and Sheffield English carrier sentences for subjects who were judged to have changed their accent less. There was no interaction with change in accent rating: these subjects did not change their best exemplar locations for *bud*, *cud*, *could* or *bath* after experience of living in a multidialectal environment.

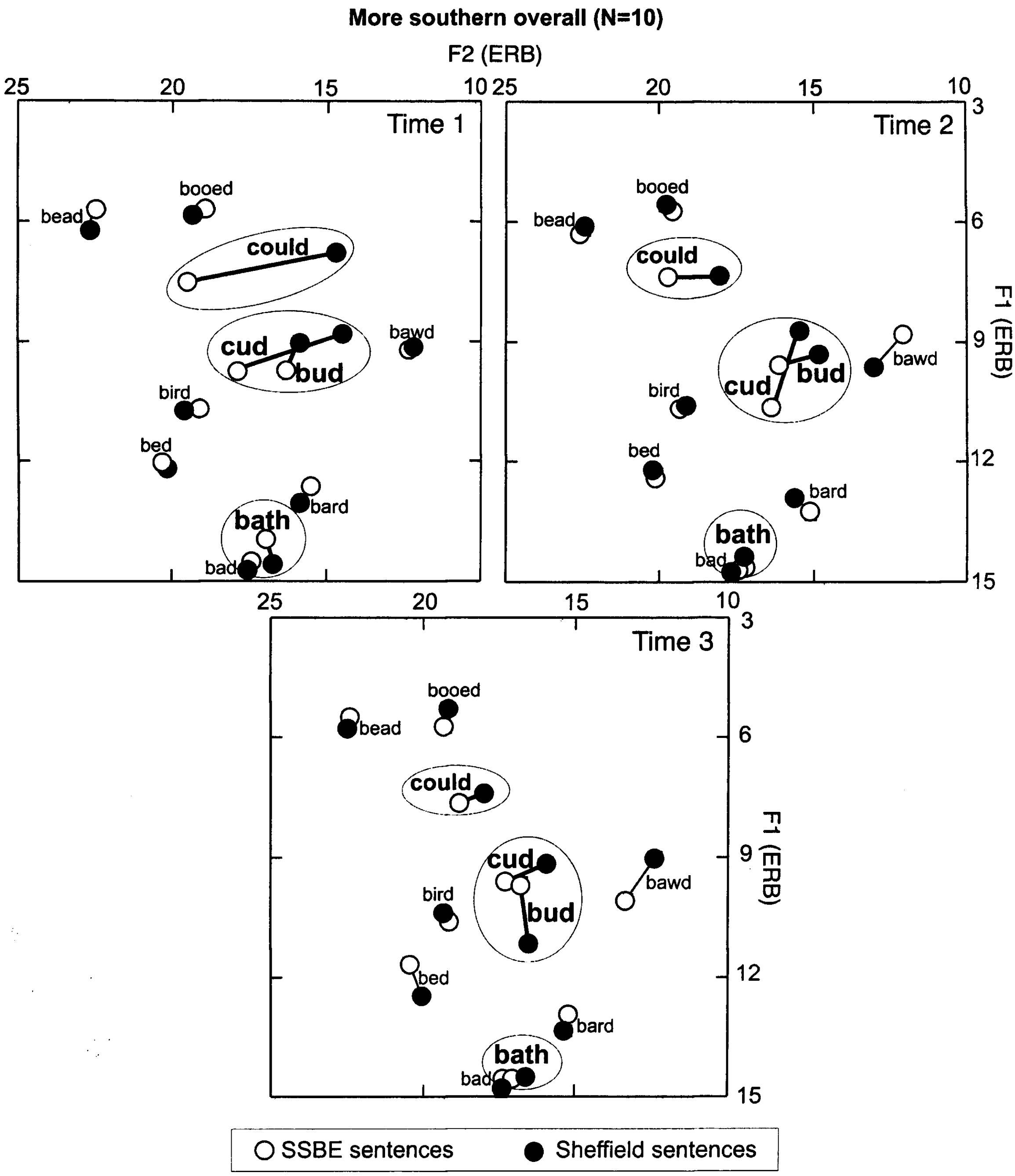

Figure 13 Average F1 and F2 formant frequencies (ERB) for best exemplars in SSBE and Sheffield English carrier sentences for subjects who were judged to have a more southern accent overall. These subjects chose significantly lower vowels (i.e., higher F1, more southern) at T1, T2, and T3 for *bud* and *cud* than did listeners who were judged to have a more northern accent overall. There were no significant interactions with overall accent rating for any other vowels.

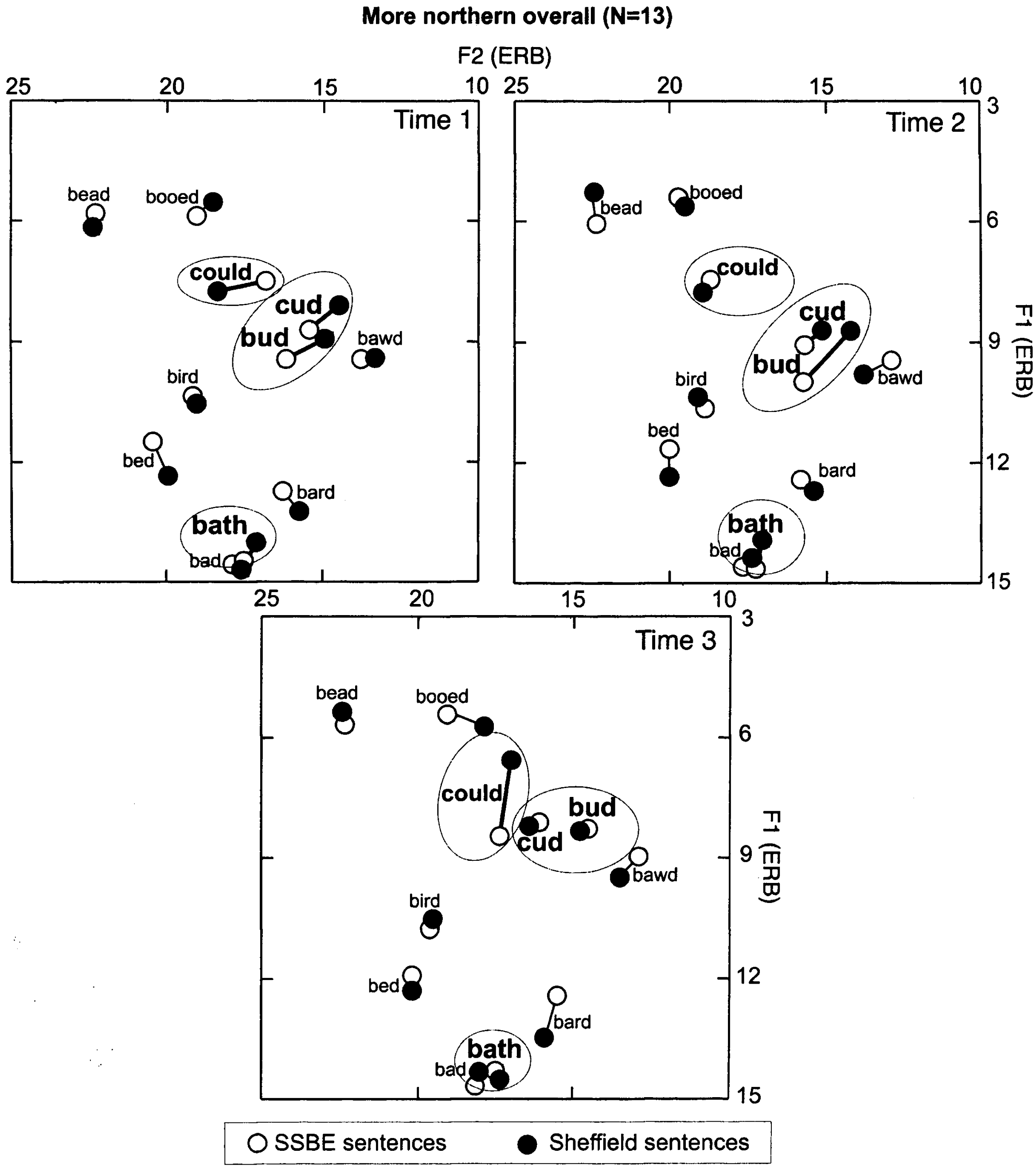

Figure 14 Average F1 and F2 formant frequencies (ERB) for best exemplars in SSBE and Sheffield English carrier sentences for subjects who were judged to have a more northern accent overall. These subjects chose significantly higher vowels (i.e., lower F1, more northern) for *bud* and *cud* at T1, T2 and T3 than subjects who were judged to have a more southern accent overall. There were no other significant interactions with overall accent rating.

Table XVI Average durations (ms) for all subjects (N=23) in SSBE and Sheffield English carrier sentences at T1, T2, and T3.

Word	SSBE			Sheffield English		
	T1	T2	T3	T1	T2	T3
Bud	72.4	70.6	75.4	74.0	64.4	72.3
Cud	59.0	70.2	71.8	59.1	68.3	62.9
Could	73.0	88.0	81.3	76.2	71.7	78.9
Bath	80.8	95.7	79.6	70.7	79.3	79.6
Bad	87.0	103.1	95.0	67.5	86.5	89.2
Bard	157.7	164.3	166.6	148.8	148.5	157.2
Bawd	155.6	145.7	145.0	140.8	144.3	145.9
Bead	125.9	135.8	128.1	118.7	136.3	130.4
Bed	70.3	84.7	74.6	60.3	67.1	62.7
Bird	125.7	131.4	151.4	116.2	144.0	134.7
Booed	152.4	132.7	146.9	123.9	146.9	135.3

There was no significant interaction with change in accent rating, $p > 0.05$, indicating that changes in subjects' best exemplar locations for these words were not linked to changes in their spoken accent. This is confirmed in Figs. 11 and 12; counter to what had been expected, subjects who changed their accent more over time (i.e., their change in accent rating was greater than or equal to the median change in accent rating for all subjects) do not appear to be choosing different best exemplars for *bud* and *cud* in SSBE and Sheffield English sentences at T1, T2 or T3 than those who changed their accent less (i.e., their change in accent rating was less than the median change in accent rating for all subjects).

However, there was a significant interaction of overall accent rating and sentence context, $F(1,17) = 4.99, p < 0.05$. This can be seen clearly in Fig. 15. Subjects who had a more southern accent (i.e., their overall accent rating was greater than or equal to the median overall accent rating for all subjects) chose different best exemplar locations in SSBE and Sheffield English carrier sentences at T1, T2 and T3 for *bud* and *cud*. In the SSBE context, they chose a lower vowel (i.e. higher F1), which does not exactly match what native speakers of that accent would produce, but matches what northerners would produce when interacting with southerners (see Chapter 2). In Sheffield English sentences, they chose a higher vowel (i.e. lower F1), which better matches what they would produce in their native accent. Subjects who were judged to have a more northern accent overall (i.e., had an overall accent rating less than the median overall accent rating for all subjects), chose a best exemplar in both SSBE and Sheffield English carrier sentences that matches what northern speakers would produce.

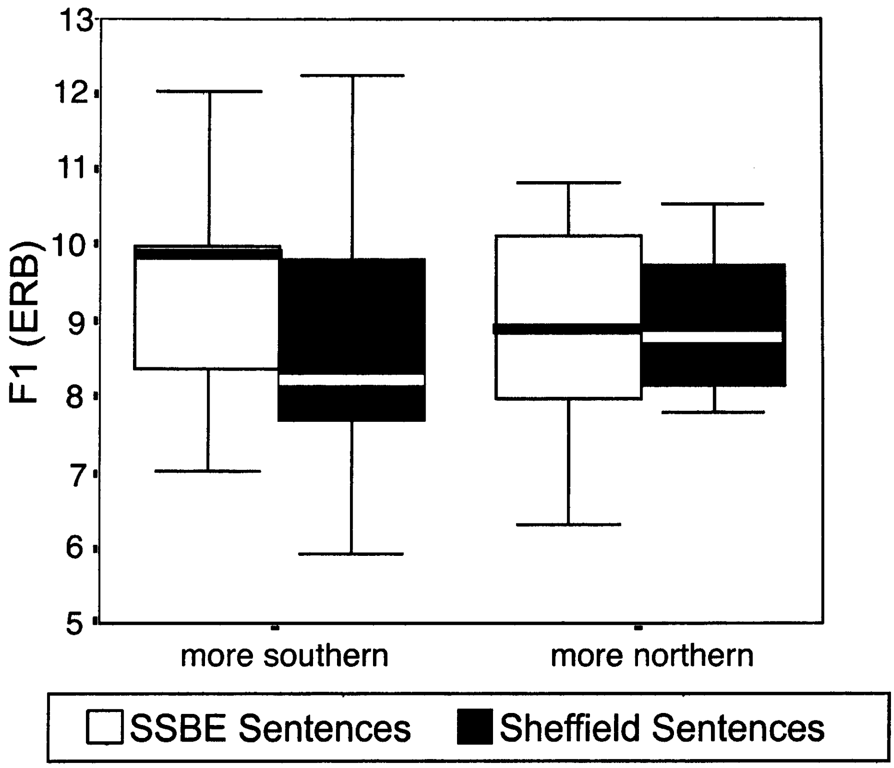

Figure 15 Boxplot of F1 formant frequencies (ERB) for *bud* in SSBE and Sheffield English carrier sentences, averaged across time, and grouped according to overall accent rating. Subjects who were judged to sound more southern normalized this vowel for accent but subjects who were judged to sound more northern did not normalize.

6.2.3.2. Could

Separate repeated measures MANOVA analyses tested whether subjects changed their vowel categorization decisions for *could*, and whether these changes were linked to the change in accent rating or overall accent rating. Time (T1, T2 and T3) and sentence context (SSBE or Sheffield English) were coded as within-subject variables, and change in accent rating and overall accent rating were coded as covariates.

For F2, there were no significant main effects or interactions, $p > 0.05$, indicating that the differences in this dimension shown in Fig. 13 in SSBE and Sheffield English sentences at T1 and T2 were unreliable. There were also no significant main effects or interactions for duration, $p > 0.05$. This confirms that subjects did not change the duration of their best exemplar for *could* as a result of experience of living in a multidialectal environment, and that they chose similar durations in SSBE and Sheffield English carrier sentences, as shown in Table XVI.

For F1, there were no main effects, but there was a significant interaction of time and change in accent rating, $F(1,17) = 14.64, p < 0.01$. The effects of this interaction can be seen clearly in Fig. 16. Subjects who changed their accent less over time did not change their best exemplar location, and chose a vowel in SSBE and Sheffield English carrier sentences that would be appropriate in their native accent. Subjects who changed their accent more over time, changed their best exemplar location in SSBE and Sheffield English contexts: they chose a lower vowel (i.e. higher F1) in SSBE and Sheffield English carrier sentences at T3 that was more similar to the vowel that southerners use when producing words like *bud* and *cud*. This change is surprising as southerners do not produce *could* with a low central vowel: they produce *could* using the same vowel as northerners, the high back vowel [ʊ].

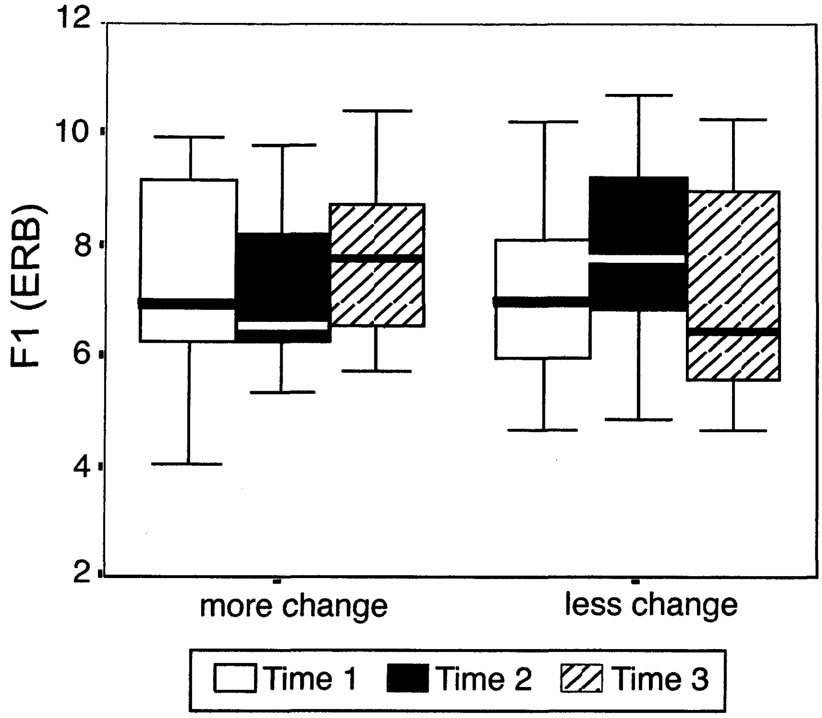

Figure 16 Boxplot of F1 formant frequencies (ERB) for *could* in SSBE and Sheffield English sentences at T1, T2, and T3, grouped according to the change in accent rating. Subjects who changed their accent more, changed their best exemplar location as a result of experience of living in a multidialectal environment; they chose a lower vowel (i.e., higher F1) at T3 than at T1. Subjects who changed their accent less did not significantly change their best exemplar location.

6.2.3.3. *Bath*

Separate repeated measures MANOVA analyses investigated whether subjects changed their vowel categorization decisions for *bath*. Time (T1, T2 and T3) and sentence context (SSBE or Sheffield English) were coded as within-subject variables, and change in accent rating and overall accent rating were coded as covariates.

The analyses demonstrated that there were no significant main effects or interactions of change in accent rating and overall accent rating for F1 or F2, $p > 0.05$, indicating that subjects did not change their perceptual categorizations for *bath*. This can be seen clearly in Figs. 10-14. As displayed in Fig. 10 all subjects chose similar best exemplar locations for *bath* in SSBE and Sheffield English carrier sentences at T1, T2 and T3, and there were no differences in subjects' best exemplar locations in SSBE and Sheffield English carrier sentences when subjects were grouped by change in accent rating (Figs. 11 and 12) or overall accent rating (Figs. 13 and 14).

There were also no significant main effects or interactions of change in accent rating and overall accent rating for duration, $p > 0.05$. This indicates that although subjects appeared to choose a longer vowel in SSBE than Sheffield English carrier sentences (Table XVI), this difference was unreliable.

6.2.3.4. *Other vowels*

Separate repeated measures MANOVA analyses tested whether there were any differences in F1, F2 and duration for all other words. Word (i.e., all words other than *bud*, *cud*, *could*, and *bath*), sentence context, and time (T1, T2, and T3) were coded as within-subject variables, and change in accent rating and overall accent rating as covariates. There was a main effect of word for F1, $F(1,14)=51.91$, $\underline{p} <$

0.001, and F2, $F(1,14)=13.69, p < 0.001$, demonstrating that different words had different formant frequency values. However, there were no main effects of sentence context or time, and no between-subject effects of change in accent rating or overall accent rating $p > 0.05$, indicating that subjects did not change their best exemplar locations for any other words.

For duration there were no significant main effects or interactions, $p > 0.05$; subjects chose similar vowel durations for each target word in both SSBE and Sheffield English carrier sentences at T1, T2, and T3. However, when the change in accent rating and overall accent rating covariates were omitted, there was a highly significant main effect of word, $F(1,21)=34.97, p < 0.001$, indicating that subjects chose different durations for different words. This can be clearly seen in Table XVI: subjects chose similar vowel durations in both sentence contexts at T1, T2 and T3.

6.3. Experiment 3: Word recognition in noise

6.3.1. Introduction

As previously explained (Chapter 5, Section 5.4), it is possible that changes in production and vowel categorization processes also affect word recognition. One could imagine that subjects who had changed their vowel categorization processes after experience with SSBE might be better able to recognize speech in noise produced in SSBE (see Eisner and McQueen, in press). This experiment investigated whether subjects' performance in word recognition tasks improved as a result of experience of living in a multidialectal environment, and whether it was linked to the amount that they changed their accent or their overall accent rating. A speech in noise paradigm was used because it was felt that this provided the best way of investigating the changes of interest in this study, i.e., whether the potential changes in phonetic categorization for words like *bud*, *cud*, and *bath* affected word recognition. As subjects are so familiar with SSBE, it was hypothesized that a testing paradigm such as a lexical decision task or phoneme spotting task would not be sensitive enough to reveal differences in word recognition. Subjects were tested in their recognition of sentences in noise produced in an SSBE and northern English accent. An additional task was given at T3 that focused more closely on vowel identification and tested subjects' online ability to switch between northern and southern English accents.

6.3.2. Method

6.3.2.1. Subjects

The subjects were the same as those in Experiment 1 in all respects.

6.3.2.2. Stimuli and Apparatus

Sentence Recognition Task. The stimuli were the BKB sentences (Bench et al., 1979). The BKB sentences were chosen as they form a standardized sentence list that is widely used as an assessment tool in clinical and non-clinical tests of speech perception. Each sentence comprises 3 highly-familiar keywords that are identified by the listener, e.g. The <u>house</u> had <u>nine</u> <u>rooms</u>, They are <u>buying</u> <u>some</u> <u>bread</u> (keywords are underlined). All 21 lists of 16 sentences each were used. The sentence lists were recorded by two female speakers of different accents; a northern English speaker and a southern English speaker. The northern English speaker had been born and raised in Ashby, and had lived there until age 18 when she had moved to Birmingham to attend university. She had lived in Birmingham for 5 years, but was able to produce an Ashby accent that was judged by the researcher (trained phonetician, originally from Ashby) to sound like that of a native speaker. The stimuli were recorded onto DAT using a Sony microphone in a sound-attenuated room. The stimuli for the SSBE speaker were taken from existing recordings made by an SSBE speaker at University College London approximately 5 years ago. These stimuli have been used widely as experimental stimuli in the Department of Phonetics and Linguistics, University College London.

Stimuli were played at a sampling rate of 16 kHz using a computer sound card. Subjects listened over headphones (Sennheiser HD 414) in a quiet room in the researcher's or subject's home.

Word recognition task. The stimuli consisted of target words in the phonetic environments /b/-V-/d/, /b/-V-/θ/, and /k/-V-/d/ produced in the carrier sentence *I'm asking you to say the word [ ] please.* As in the goodness optimization task, the /b/-V-/θ/ words were included because northerners and southerners produce *bath* with different vowels. The /k/-V-/d/ words were included because northerners do not have the southern vowel [ʌ] in their native accent; instead northerners produce *could* and *cud* with the same vowel, [ʊ], such that *cud* and *could* are homophones.

The stimuli were produced by three male speakers of northern and southern English accents. One speaker was a northern English speaker from Leicestershire, who had been born and brought up in Leicestershire until age 18 years. The speaker had moved to London to attend university where he had lived for 7 years, but was still able to produce a Leicestershire accent that was judged by the researcher to sound like that of a native speaker. The second speaker was an SSBE speaker, who had lived in the south of England since age 2 years. The third speaker was the speaker who produced the stimuli described in Study 1 for the Goodness Optimization task (Chapter 4) and Study 2, Experiment 2 (Chapter 6, Section 6.2). He recorded stimuli in both northern (Sheffield) and SSBE accents (see Chapter 4, Section 4.2.1.2). This gave four sets of stimuli; two sets produced in northern English, and two sets produced in SSBE. Each speaker recorded multiple repetitions of each experimental word in the carrier sentence and one example of each was selected for use in the experiment.

The stimuli were recorded onto DAT using a Sony microphone. Recordings were made in a quiet room at the researcher's or speaker's home. The stimuli were transferred to computer at a sampling rate of 16 kHz and were mixed with continuous talker babble at a signal-to-noise ratio (SNR) of −12 dB using Praat (Boersma and Weenink, 2004). The talker babble was from the Institute of Hearing Research (S. Rosen, personal communication), and has been widely used in word recognition experiments. The SNR was selected based on preliminary results from the sentence recognition task and pilot experiments conducted at University College London.

Stimuli were played at a sampling rate of 16 kHz using a computer sound card. Subjects listened over headphones (Sennheiser HD 414) in a quiet room in the researcher's or subject's home.

6.3.2.3. *Procedure*

Sentence recognition task. There were six blocks of testing, three for each speaker. Three blocks were completed on the first day of testing and three on the second day. The blocks alternated between the two speakers (i.e., subjects did not complete two consecutive blocks for the same speaker), and the order of presentation was counterbalanced across subjects. Subjects were randomly assigned a set of 10 lists selected from lists 2-21 (i.e., 160 sentences; maximum number of trials per block is 20). List 1 was used for familiarization. Subjects were assigned a different set of lists at T1, T2, and T3.

The three blocks of testing took approximately 5 minutes to complete. Before beginning the first block of testing, subjects were familiarized with the procedure using List 1. Subjects then completed the three test blocks. Subjects heard the sentence and then repeated back to the researcher what they thought they had heard. If they were unsure what they had heard, subjects were instructed to guess. If subjects were unable to respond they were reassured that this was normal and indeed inevitable due to the test design. Subjects heard each stimulus only once. The researcher scored the number of keywords (out of three) the subject repeated. For example, if the sentence "The house had nine rooms" was presented, then the listener had to repeat the words "house", "nine" and "rooms" in order to score 3 points. If the listener only repeated "house" and "nine" then he/she would only score 2 points.

The task used an adaptive procedure to find subjects' noise thresholds. The procedure tracks the amount of noise a listener can tolerate in a speech signal before word identification is impaired. The stimuli were presented at varying SNRs, and were put together on-line; noise was presented at a fixed level of 71 dBA, and an adaptive procedure varied the SNR of the stimulus and noise. The masker noise was a broadband noise (0-5000 Hz), pre-synthesized and recorded in Praat (Boersma and Weenink, 2004).

The adaptive procedure used a modified Levitt procedure (Levitt, 1971; Baker and Rosen, 2001). In the standardized Levitt procedure (Levitt, 1971) the SNR of the initial stimulus is set 'above' the identification threshold (i.e., the stimulus is easily identified). The SNR of subsequent presentations is governed by the step-size used and the response to the stimulus. The step-size used is 2 dB and the test starts with a presentation at an SNR of +10 dB. If a listener responds correctly to the stimulus presented at +10 dB then the SNR is lowered by 2 dB to +8 dB. The SNR continues to be lowered in steps of 2 dB until the listener responds incorrectly when it is raised by 2 dB. The first incorrect response is known as the first reversal and the threshold is found after a further 8 reversals have been completed.

The modified Levitt procedure was used in this experiment because it is able to estimate the noise threshold more efficiently than the standard Levitt procedure. The procedure starts with an easy stimulus with an SNR of +10 dB (i.e., above threshold) in order to enable subjects to tune-in to the talker. The SNR then decreases in 8 dB steps after each correct response (i.e., becomes more difficult), until the first reversal (i.e., an incorrect response). After the first reversal the standard Levitt procedure is used; the SNR is changed in steps of 2 dB for a further 8 reversals. If the listener responds correctly then the SNR decreases by 2 dB (i.e., becomes more difficult) and if the listener responds incorrectly the SNR increases by 2 dB (i.e., becomes easier). This procedure is more efficient because it enables the point at which the listener cannot identify the stimulus to be found more quickly.

Sentences are scored as correct when subjects repeat all three keywords. If subjects only repeat 1 or none of the keywords, then the sentence is scored as incorrect. If subjects repeat 2 keywords then the SNR remains the same and this is not counted as a reversal. As such the procedure can be described as tracking the 66.6% correct point on the psychometric function, (i.e., the SNR is not changed with a response of 2 out of 3, or 66.6%, keywords correct). The test terminated when subjects completed 8 reversals or after 20 stimuli had been presented.

Word identification task. There was a single testing session consisting of 320 trials. Stimuli were presented in a randomized order and subjects heard each stimulus presentation only once. The experiment was self-paced and lasted approximately 25 minutes. At the start of each session, subjects completed 5 practice trials to familiarize them with the format of the experiment. They then made 320 identifications; 4 repetitions of 16 experimental words for each of the 4 speakers. The experimental words were *bad, bard, bed, bird, bud, bod, bawd, bid, bead, booed, cud, could, cooed, Beth, birth,* and *bath*. Subjects gave their response by pointing and clicking with a computer mouse on a button containing the word they thought they had heard. Although there were only 16 experimental words, subjects selected their responses from 23 words that covered the entire range of possible CVC words for each of the three phonetic environments. This was to prevent listeners from using consonantal context as a clue to word identification when identifying the /b/-V-/θ/ and /k/-V-/d/ stimuli. Subjects were instructed to click on the button containing the word they thought best represented what they had heard; they were not told that there were a smaller number of experimental words than responses.

6.3.3. Results

Sentence Recognition Task. As displayed in Fig. 17, all subjects performed better with SSBE speech. Some subjects also appeared to improve with SSBE after experience of living in a multidialectal environment (Table XVII). A repeated measures MANOVA tested whether the observed differences in performance were linked to the change in accent rating and the overall accent rating. Speaker (northern or SSBE) and time (T1, T2 and T3) were coded as within-subject variables, and change in accent rating and overall accent rating were coded as covariates.

There was a main effect of speaker $F(1,18) = 75.15, p < 0.001$, confirming that all subjects performed better with the SSBE speaker. It is plausible that this is because of differences in speaker intelligibility. The northern speaker was from Ashby, and one feature of the local accent is that it has a flat intonation contour. In contrast, the SSBE speaker had very a marked intonation pattern. This difference, combined with the fact that subjects' were highly familiar with SSBE through the media (Foulkes and Docherty, 1999) may have resulted in the SSBE speaker being more intelligible in noise than the northern speaker. There was no main effect of time and no interaction of time and speaker, $p > 0.05$.

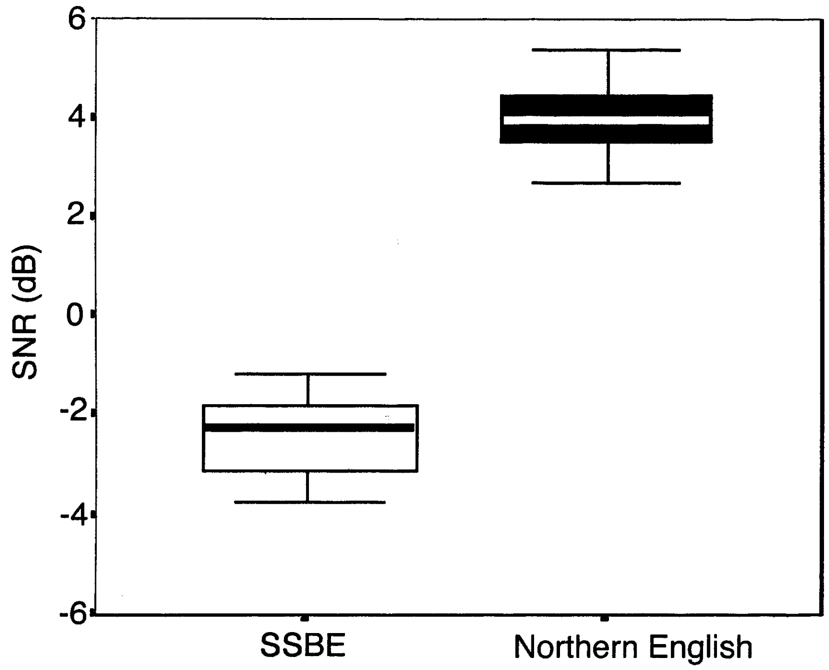

Figure 17 Boxplot of subjects' average SNR (signal-to-noise ratio) threshold for the SSBE and northern English speaker, averaged over T1, T2 and T3.

Table XVII Average SNR thresholds for all subjects (N=23) at T1, T2, and T3, in SSBE and northern English accents in the sentence recognition task.

Subject	SSBE			Northern English		
	T1	T2	T3	T1	T2	T3
F01	-0.58	-0.37	-4.70	8.40	3.83	4.47
F02	-2.94	-4.61	-3.61	3.91	3.37	3.24
F03	-2.21	-2.49	-3.36	6.16	4.34	1.38
F04	-4.53	-1.92	-3.45	4.61	4.90	1.50
F05	-4.60	-2.57	-4.81	3.62	3.59	1.62
F06	-5.27	-2.45	-4.81	2.08	2.61	3.28
F07	-4.11	-2.96		1.90	3.52	
F08	-2.63	-2.39	-3.82	4.33	4.85	3.29
F09	-2.59	-3.01	-3.87	2.56	2.85	1.77
F10	-4.48	-2.16	-2.64	2.76	6.69	2.66
F11	-4.39	-3.43	-2.61	3.62	3.27	3.37
F12	-3.94	-4.73	-4.46	3.89	3.15	1.76
F13	-2.95	-3.38	-3.02	3.46	2.74	2.20
F14	-2.37	-4.33	-4.81	3.14	5.32	2.09
F15	-3.23	-4.22	-5.26	3.74	3.58	3.67
F16	-3.64	-2.78	-5.33	4.02	2.49	3.98
M01	1.50	-4.49	-4.30	1.57	4.44	3.84
M02	-3.65	-2.13	-3.23	1.99	5.14	2.50
M03	-2.95	-4.61	-5.89	1.54	2.48	2.75
M04	-4.93	-3.65		3.12	4.22	
M05	-1.42	-3.17	-5.62	1.78	3.82	1.13
M06	-3.31	-4.87	-5.07	0.54	2.59	4.45
M07	-3.08	-3.96	-6.58	5.16	4.77	1.36

There were no interactions between the change in accent rating and speaker, $p > 0.05$, suggesting that subjects who were judged to have changed their accent had no advantage over subjects who had not changed their accent. However, there was a significant interaction between speaker and overall accent rating, $F(1,18) = 4.80, p < 0.05$. This can be seen in Fig. 18. Subjects who had a more southern accent performed better with SSBE speech than those who had a more northern accent, and subjects who had a more northern accent performed better with northern speech.

Word identification task. As displayed in Table XVIII, 15 out of 23 subjects performed better on average with SSBE speakers than with northern English speakers. There were also individual differences in the results: some subjects appeared to perform better than others with both SSBE and northern English speakers. These differences appeared to occur predominantly in the identification of *bud*, *cud*, and *could*. There were few differences for *bath* and all other words. The potential differences in performance for *bud*, *cud* and *could*[5], *bath*, and all other vowels were tested in separate repeated measures MANOVA analyses. In each MANOVA, word (*bud*, *cud*, and *could, bath* or all other vowels) and stimulus accent (SSBE or northern English) were coded as within-subject variables. The change in accent rating and overall accent rating were coded as covariates.

For *bath* and all other vowels there were no main effects or interactions, $p < 0.05$, confirming that there were no differences in performance for these words between subjects or in the different accent conditions.

---

[5] *Bud*, *cud*, and *could* were tested in the same MANOVA because *cud* and *could* are produced as homophones or near-homophones in northern English accents.

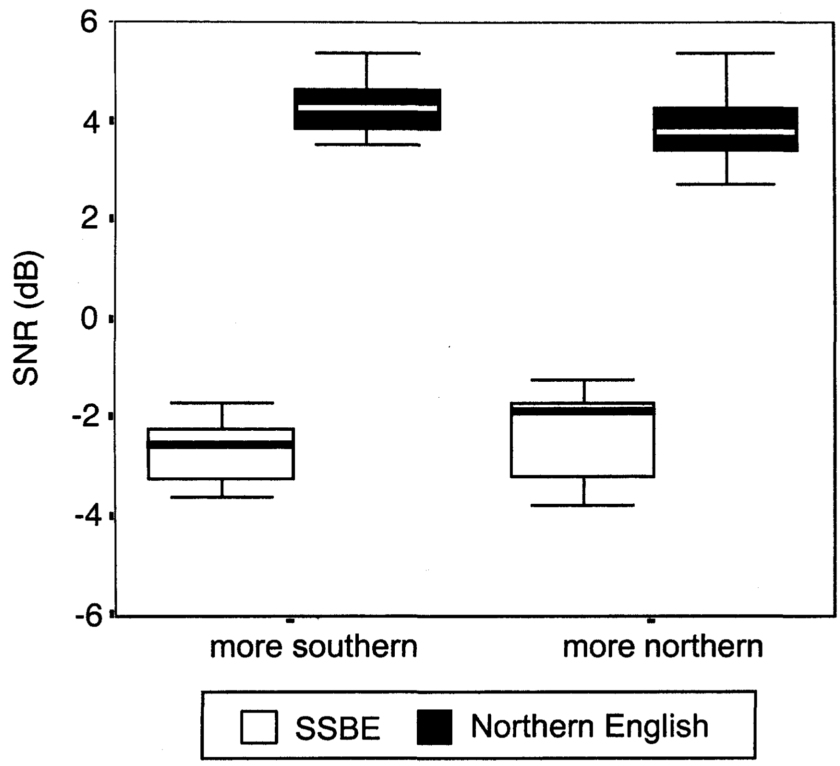

Figure 18 Boxplot of subjects' average SNR (signal-to-noise ratio) threshold for the SSBE and Northern English speaker, averaged over T1, T2, and T3, and grouped according to overall accent rating. Subjects who sounded more southern overall performed better with the SSBE speaker than those who sounded more northern overall. Subjects who sounded more northern overall performed better with the northern English speaker than did those who sounded more southern overall.

Table XVIII Percentage identification scores for *bud*, *cud*, *could*, *bath*, and all other words for each subject in the word identification task.

Subject	Bud		Cud		Could		Bath		Other words		Average	
	SSBE	NE	SSBE	NE	SSBE	NE	SSBE	NE	SSBE	NE	SSBE	NE
F01	100	80	20	0	60	100	100	90	65.0	61.7	69.0	66.3
F02	0	80	0	0	80	90	100	90	63.3	72.5	48.7	66.5
F03	90	90	50	60	50	30	90	100	64.2	69.2	68.8	69.8
F04	60	40	50	50	70	30	100	100	61.7	68.3	68.3	57.7
F05	80	60	60	70	60	50	80	100	64.2	65.0	68.8	69.0
F06	50	100	50	40	90	40	100	100	79.2	72.5	73.8	70.5
F07	100	70	60	20	70	80	100	100	78.3	75.8	81.7	69.2
F08	80	90	30	30	70	30	100	100	70.0	57.5	70.0	61.5
F09	70	50	70	40	60	30	100	100	72.5	69.2	74.5	57.8
F10	30	60	60	70	40	20	90	100	59.2	64.2	55.8	62.8
F11	70	80	70	30	70	50	100	100	69.2	68.3	75.8	65.7
F12	50	10	80	30	100	80	100	70	70.0	70.0	80.0	52.0
F13	70	80	0	0	100	90	100	100	62.5	66.7	66.5	67.3
F14	60	80	90	50	80	60	100	100	77.5	78.3	81.5	73.7
F15	50	0	60	40	10	10	90	100	62.5	72.5	54.5	44.5
F16	30	10	80	20	90	40	100	100	76.7	76.7	75.3	49.3
M01	90	40	80	10	100	100	100	100	75.0	71.7	89.0	64.3
M02	60	40	80	0	90	80	100	80	61.7	57.5	78.3	51.5
M03	40	50	80	20	90	50	100	100	81.7	73.3	78.3	58.7
M04	30	60	60	70	40	20	90	100	59.2	64.2	55.8	62.8
M05	30	80	60	50	40	50	100	100	70.0	75.0	60.0	71.0
M06	0	100	70	10	90	40	100	100	86.7	77.5	69.3	65.5
M07	70	80	0	0	100	90	100	100	62.5	66.7	66.5	67.3
Average	57	62.2	55	30.9	71.74	54.78	97.4	97	69.2	69.3		

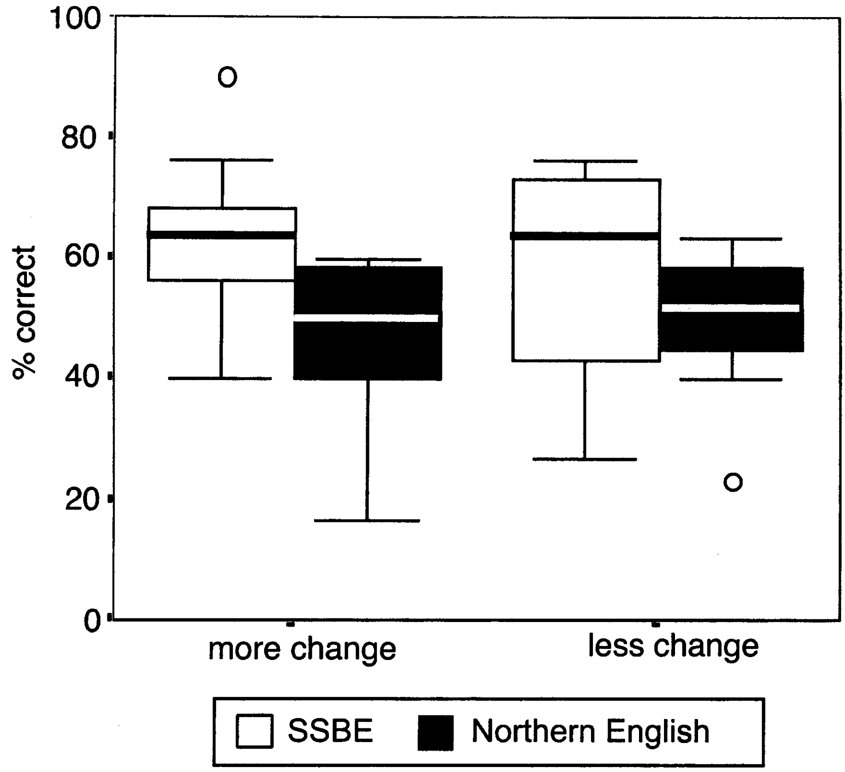

Figure 19 Boxplot of subjects' percentage correct identifications for *bud*, *cud*, and *could* in SSBE and northern English accents, grouped according to change in accent rating. Subjects who changed their accent more performed significantly better with the SSBE speaker than those who did not change their accent.

For *bud*, *cud*, and *could* there were no significant main effects of word or accent, $p < 0.05$. There was a significant interaction of change in accent rating and accent, $F(1,18) = 4.39, p = 0.05$, indicating that subjects who changed their accent to sound more southern improved in their identification of these words in SSBE. However, this is a very small effect (Fig. 19). Although subjects who changed their accent less show more variability in their responses in the SSBE condition and more overlap with their performance in the northern English condition, subjects who changed their accent do not appear to have any clear advantage in the SSBE condition.

There was also a significant three-way interaction of word, stimulus accent and overall accent rating, $F(1,18) = 9.28, p < 0.01$, indicating that subjects who sounded more southern overall were better at identifying these words in SSBE. Fig. 20 suggests that this interaction is being driven by differences in performance for *bud*. Both groups of subjects performed similarly with *cud* and *could* in the SSBE and northern English conditions, but for *bud*, subjects who had a more southern accent performed better with SSBE speakers than did those with a more northern accent. In contrast, subjects who had a more northern accent performed better than subjects with a more southern accent in the northern English condition.

These differences were tested in separate repeated measures MANOVA analyses, with stimulus accent (SSBE or Sheffield English) coded as a within-subject variable, and the overall accent rating coded as a covariate. For *cud* and *could* there was no significant main effect or interaction with overall accent rating, $p > 0.05$, confirming that both groups of subjects performed similarly in both the SSBE and northern English conditions for these words. For *bud* there was a main effect of accent, $F(1,21) = 5.44, p < 0.05$. This can be clearly seen in Fig. 21; all subjects were better at identifying *bud* in a northern English accent than in SSBE. There was also an interaction with stimulus accent and overall accent rating, $F(1,21) = 6.42, p < 0.05$, confirming that subjects who produced *bud* with a more southern vowel were better able to identify this word in the SSBE condition than those who produced this word with a more northern vowel. It is unclear why listeners were better able to identify *bud* than *could* and *cud*.

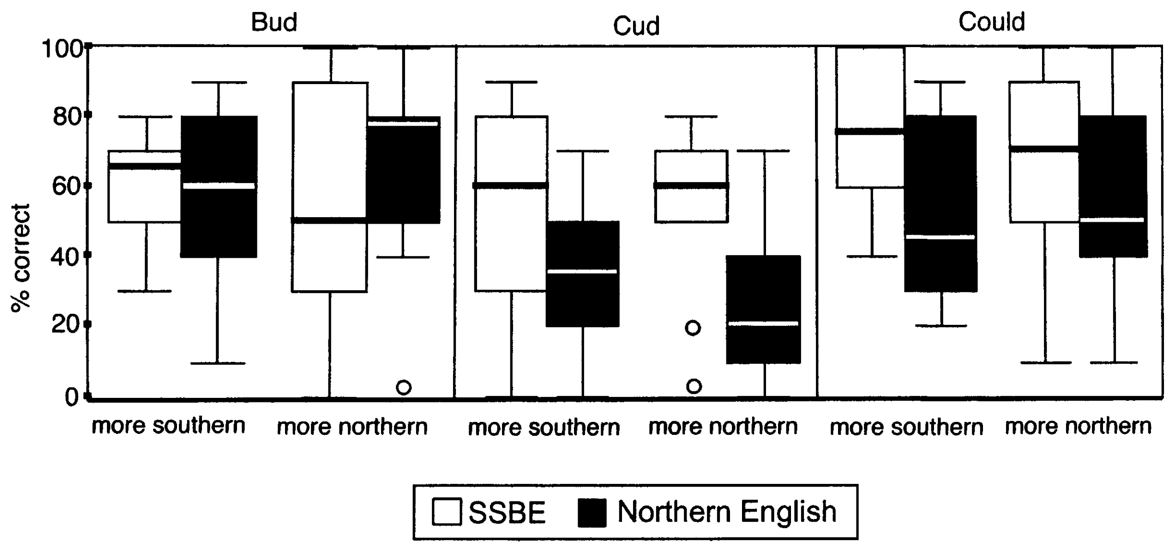

Figure 20 Boxplot of subjects' percentage correct identifications for *bud*, *cud*, and *could* in SSBE and northern English accents, grouped according to subjects' overall accent rating. Subjects who sounded more southern overall were significantly better at identifying *bud* in an SSBE accent than those who sounded more northern. Subjects who sounded more northern performed better with *bud* in a northern English accent than subjects who sounded more southern.

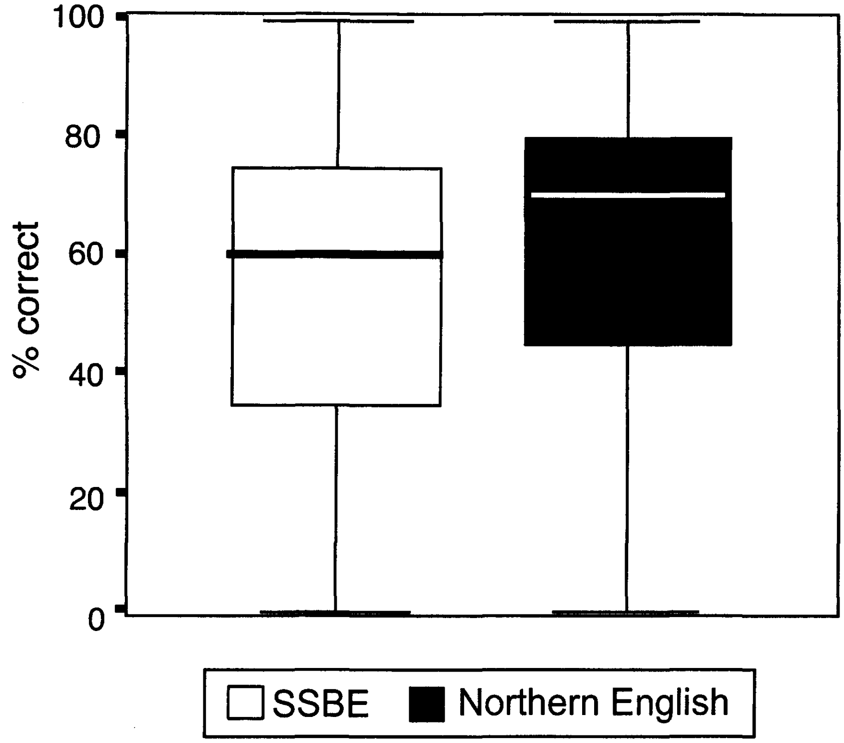

Figure 21 Boxplot of subjects' percentage correct identifications for *bud* in SSBE and northern English accents. All subjects were better at identifying *bud* in northern English sentences than SSBE sentences.

6.4. Discussion

6.4.1. Change in production and perception

Study 1 demonstrated that subjects who had the experience of living in a multidialectal environment normalized the vowels in *bud* and *cud* – but not *bath* – for southern and northern English accents. The patterns of normalization reflected differences in subjects' linguistic experience. When individuals living in London heard sentences that were similar to their native accent, they chose formant frequencies for *bud* and *cud* that matched what speakers of that accent would produce. When they heard sentences that did not match their native accent, they chose centralized vowels for *bud* and *cud* that matched how they would produce these words when interacting with a native speaker of that accent. Subjects who were less experienced with a non-native accent (i.e., Ashby subjects) did not normalize for accent at all; they chose vowels in SSBE and Sheffield English carrier sentences that match what they would produce in their native accent. It was hypothesized that these subjects did not normalize for accent because they had not had the experience of changing their production when interacting with SSBE speakers.

Based on this evidence it was hypothesized that the subjects in Study 2, born and raised in the north of England and with no previous experience of regularly interacting with SSBE speakers, would change their production and perception after experience of living in a multidialectal environment. It was hypothesized that subjects would change their productions of the vowels in words like *bud* and *cud* to better fit in with SSBE speakers, and that this would affect their perceptual processes, such that they would choose different best exemplar locations for *bud* and *cud* in SSBE and Sheffield English carrier sentences, and would be better able to recognize SSBE speech.

The results demonstrated that subjects did change their accent as a result of experience of living in a multidialectal environment. Of the 23 subjects tested, 19 were judged to have changed their accent. An acoustic analysis demonstrated that subjects changed their production of the vowels in words like *bud* and *cud*. This shift occurred in both the F1 and F2 dimension: subjects produced a more fronted and lower vowel in these words at T3 than at T1. However, the change in both the accent ratings and production was often small and there was a lot of variability in the data:

Some subjects changed their accent more than others and some subjects were judged to sound more southern than others at both T1 and T3.

Data from sociolinguistic interviews revealed that there was a relationship between the amount subjects changed their accent and their friendship groups at university. Subjects who changed their accent the most had friendship groups formed predominantly of SSBE speakers. They reported that they had changed their accent to better fit in with their friendship groups and because they wanted to be identified with this community. It is plausible then, that the subjects who changed their accent did so because they had more experience of interacting with SSBE speakers and were highly motivated to fit in with this community.

The acoustic analysis also revealed that subjects did not change all aspects of their accent to better fit in with SSBE speakers. They changed the vowels in *bud* and *cud*, but in words like *bath* they retained their native pronunciation. Sociolinguistic factors may also be able to explain these differences. Subjects, who changed their accent to better fit in with their friendship groups, also felt that it was important to retain some aspects of their regional identity. It is possible that they retained their native pronunciation in words like *bath* in order to continue signalling loyalty to their native community, but changed their production of the vowel in words like *bud* and *cud* in order to identify themselves with their new multidialectal community.

It has been suggested that it is important that the features that are changed when adapting to a non-native accent do not signal any other particularly well-defined variety, because of the potential signalling of disloyalty to local norms (e.g., Foulkes and Docherty, 1999). This is thought to be particularly important where the standard accent is concerned, due to connotations of 'snobbishness' (Hickey, 1999). It is interesting then, that although subjects who had changed their accent to sound more southern felt that they had retained certain regional variants in order to continue signalling loyalty to their native accent community, they were often regarded as 'sounding posh' by other northerners who had not changed their accent. For example, M06, a good friend of F15 at school and who attended the same university commented that unlike F15, M06 "hung around with all the posh horsey people and now she sounds all la-di-dah". Both F15 and M06 had a strong northern accent at T1 (both were given an average rating of above 7), but at T3 F15 was given a much lower rating than M06; F15 was given an average rating of 5.6 and M06 was given an average rating of 8.4, where a rating of 10 means that the speaker had a very strong northern accent. This suggests that whilst subjects who have changed their accent feel that they have retained certain regional variants in order to signal belonging to their native accent community, speakers of the native accent do not necessarily regard them as members of that community.

However, the changes subjects' made in production did not match the changes they made in perception. Subjects who changed their accent changed their best exemplar location for *could* – but not *bud* and *cud* – in SSBE carrier sentences. Even before experience of interacting with SSBE speakers (i.e., T1), subjects chose a relatively central best exemplar for *bud* and *cud* in the Goodness Optimization Task. After experience of interacting with SSBE speakers, subjects chose a more central vowel with a higher F1 (i.e., lower in the vowel space) in SSBE carrier sentences for *could* that was closer to the vowel they chose for *bud* and *cud* in SSBE carrier sentences. This is particularly surprising, as SSBE speakers do not produce *could* with this vowel; they produce words like *could* using the same vowel as northern English speakers, the high back vowel [ʊ]. It is possible that they chose the same best exemplar for *could* after experience with SSBE, because they were hypercorrecting (Wells, 1982b; see Section 2.3). That is, subjects were using this central vowel in all words where they would use the lower, back vowel [ʊ] in their native accent, even if this did not correspond to how southerners produce these words.

There was also little evidence to suggest that subjects were learning to normalize. Based on Study 1, it was expected that after experience of interacting with SSBE speakers, subjects would adjust their spoken accent and choose different best exemplar locations for the vowels in *bud* and *cud* in SSBE and Sheffield English carrier sentences. However, subjects who changed their accent did not normalize; although subjects chose a different best exemplar location for *could* at T3, they chose similar vowels in SSBE and Sheffield English carrier sentences.

It is unclear why subjects did not normalize. It is possible that the ability to normalize for accent takes a longer time to develop than the time-course of this study. Of the northern subjects tested in Study 1 who did normalize, all had lived in London for at least 1 year and with a mean of 8.6 years. It is thus plausible that were these subjects to be tested at a later stage evidence for accent normalization would be found. Another possibility is that subjects had learned to normalize but that the testing environment was not conducive to this being shown. The northerners tested in Study 1 who normalized for accent, lived and completed the tests in London, a multidialectal community. Due to organizational constraints, the subjects in this study were tested in Ashby, a northern English accent community. It is possible that if they had been tested whilst at university, a multidialectal environment where SSBE is regularly used, normalization effects would have been found.

Changes in production did not clearly affect word recognition processes either. Subjects who changed their accent to fit in with SSBE speakers did not show any advantage for the SSBE speakers in the sentence recognition task. There was a possible advantage in the word identification task; subjects who changed their accent appeared to be better at identifying *bud*, *cud* and *could* in SSBE speech than subjects who had not changed their accent.

6.4.2. Individual differences in production and perception

Although subjects were from similar backgrounds and had all lived in Ashby since age 5, there were differences in their accents. Some subjects were judged to sound more southern at T1, T2 and T3, and others were judged to have a particularly strong northern accent overall. An acoustic analysis of the production data showed that subjects who had a more southern accent overall produced the vowels in words like *bud* with a more central vowel at T1, T2 and T3 that was closer to how southerners produce these vowels. These subjects produced this vowel with a higher F1 (i.e., lower in the vowel space) and a higher F2 (i.e., more fronted).

Individual differences in the production data also corresponded with differences in the perceptual data. Subjects who produced more southern vowels overall in words like *bud* also chose more southern vowels for *bud* and *cud* in SSBE sentences in the Goodness Optimization Task (Experiment 2) at T1, T2 and T3. Subjects who produced more northern vowels chose more northern vowels for *bud* and *cud* in SSBE sentences. Subjects who sounded more southern overall also normalized the vowel in *bud* more. However, there was no change in the amount of normalization (i.e., these subjects did not normalize more after experience of interacting with SSBE speakers).

Similarly, differences in word recognition processes were linked to differences in production. Subjects who produced more southern vowels performed better with SSBE speech in the sentence recognition task and the word identification task than did those subjects who produced more northern vowels at T1, T2 and T3. In the word recognition task, subjects who produced *bud* with a more southern vowel were better able to identify this word in noise than subjects who produced this word with the northern vowel [ʊ].

6.4.3. Summary

The results from Study 2 only partially support the hypothesis from Study 1 that subjects born and raised in the north of England and with no previous experience of regularly interacting with SSBE speakers, would change their speech production and perception after experience of living in a multidialectal environment. Specifically, it was hypothesized that subjects would change their spoken accent to better match SSBE speakers and that they would change their best exemplar locations for SSBE accented speech in a vowel categorization task, the Goodness Optimization Task (Iverson and Evans, 2003). It was also hypothesized that potential changes in production and vowel categorization might lead to changes in subjects' ability to identify SSBE speech in degraded listening conditions (i.e., speech in noise).

Some subjects did change their accent, but these changes were only small. As expected, the changes were in the vowels in words like *bud*, which are produced differently in northern and southern English accents. Adaptation to a non-native accent also appeared to be linked to sociolinguistic factors. Subjects who changed their accent had friendship groups made up predominantly of SSBE speakers and were highly motivated to change their accent to better fit in with these groups. However, they also felt that it was important to maintain some aspects of their regional identity. This was demonstrated through their retention of short [a] in words like *bath* and *grass* in both their production and perception.

There was some evidence for a link between production and perception. Subjects who sounded more southern overall, chose vowels for *bud*, *cud*, and *could* in the Goodness Optimization Task at T1, T2, and T3 that were more similar to how southerners produce these vowels (see Chapter 4, Section 4.4). However, the degree of change in production did not correspond with changes in perception; subjects changed their production of the vowel in words like *bud*, but changed their best exemplar location for *could*. This was surprising as *could* is produced using the same vowel, [ʊ], in both northern and southern English accents. It was hypothesized that this was because subjects were hypercorrecting (Wells, 1982b). Subjects who changed their accent did not learn to normalize for accent either; subjects did not choose different best exemplar locations for *bud* and *cud* in SSBE carrier sentences.

The changes in production were also greater than the changes in perception. This suggests that production of non-native speech sounds is easier and/or faster to learn than the perception of the same sounds. This is supported by evidence from L2 speech research. Flege (2003) demonstrated that even though late bilinguals may be judged to produce a language like a native speaker, they perform differently from native speakers in perceptual tasks. Bohn and Flege (1997) also found that experience with an L2 had a greater effect on production than perception. Experienced German learners of English were able to produce a new vowel contrast, the English contrast /ɛ/-/æ/, that was judged to sound like those of native speakers, but they did not rely on the same acoustic cues when discriminating between the two sounds; unlike native English listeners, they relied more on duration than spectral differences when discriminating /ɛ/ and /æ/. It is thus plausible that after a longer period of time and more experience of living in a multidialectal environment, subjects in this study may show a greater amount of change in their best exemplar locations and a greater improvement in their ability to recognize SSBE speech in noise.

7. General Discussion

7.1. Summary

Study 1 demonstrated that some subjects adjusted their vowel categorization decisions when listening to speech produced in different regional accents. Subjects living in London chose best exemplars for *bud* and *cud* – but not *bath* – in an SSBE context that differed from those they chose in a northern English context, with the patterns of variation reflecting each listener's linguistic experience. When individuals living in London heard sentences that were similar to their native accent, they chose formant frequencies for *bud* and *cud* that matched what speakers of that accent would produce. When individuals living in London heard sentences that did not match their native accent (e.g., northerners listening to SSBE speech), they chose centralized vowels for *bud* and *cud* rather than the [ʌ] and [ʊ] vowels that would normally be produced in SSBE and Sheffield accents, respectively. These patterns of variation corresponded closely with the changes in production that speakers tend to make when they live in a multidialectal environment, i.e., when they are accommodating to SSBE speakers (Wells, 1982b; Trudgill, 1986). Northerners who live in the south of England typically modify some aspects of their accent in order to fit in with southerners; they change their production of the vowel in words like *bud* and *cud* so that it becomes more centralized. However, northerners retain some aspects of their regional identity; they retain [a] when producing words like *bath* (see Section 4.4 for a full discussion of this point).

It was hypothesized that production might also explain why Ashby subjects did not perceptually normalize for accent. Ashby subjects in Study 1 chose vowels in SSBE and Sheffield English carrier sentences that were similar to their native northern vowels. It is possible that these subjects did not normalize for accent because even though they were highly familiar with SSBE through the media (e.g., Foulkes and Docherty, 1999), they had not had experience of modifying their own speech in order to fit in with SSBE speakers.

Study 2 investigated whether subjects from Ashby with no previous experience of living in a multidialectal environment, changed their production when attending university, and whether these changes were linked to perceptual changes in vowel categorization and word recognition. The results demonstrated that subjects changed their production, although the differences were small and there was a lot of variability in the results. There was some evidence for a link between production and perception. In the Goodness Optimization Task (Experiment 2), subjects chose similar vowels to those that they produced; subjects who produced more southern vowels for *bud* and *cud* chose more southern best exemplars for these words. However, the changes in production and perception occurred in different words; subjects changed their production of the vowel in *bud* and *cud*, but changed their best exemplar location for *could*. There was no evidence for normalization; although subjects changed their best exemplar locations for *could* they chose similar vowels in SSBE and Sheffield English carrier sentences.

Word recognition was similarly affected. Subjects who produced more southern vowels overall were better at identifying SSBE speech. However, there was no strong link with changes in production; subjects who changed their accent and vowel categorization processes did not show any clear change in their performance in the word recognition tasks with SSBE speech.

7.2. Comparison with existing models

7.2.1. Exemplar models of speech perception

Episodic memory research has shown that individuals store phonetically detailed representations of spoken words in long-term memory (e.g., Goldinger, 1998; Nygaard and Pisoni, 1998; Palmeri et al., 1993). It has been suggested that these exemplar representations can produce talker normalization effects if subjects compare the words that they hear to stored exemplars of speech produced by similar talkers (Johnson, 1997). Based on this research, it was hypothesized that listeners would be able to easily adapt their perceptual representations and normalize fully for accent, provided that they had experience with that accent. However, the results demonstrated that even though there were some changes in subjects' vowel categorization processes as a result of experience of living in a multidialectal environment, subjects were unable to easily adapt their vowel categorization decisions to match talkers of that accent. Both the northern subjects living in London tested in Study 1 and subjects in Study 2 chose best exemplar locations for words like *bud*, *cud* and *bath* in SSBE carrier sentences that did not match what native speakers of these accents produce, even though they were all highly familiar with SSBE.

It could be hypothesized that this was a result of the test words that were used in the Goodness Optimization Task. One could imagine that northerners chose vowels for *bud*, *cud*, and *bath* that did not match how southerners produce these words, because they did not have memories of how SSBE speakers produce these words. Although *bud* and *cud* are low frequency words, it is unlikely that northerners were unaware of how southerners produce the vowel in *bath*. *Bath* is a high frequency word, and northern English speakers are acutely aware of how this word differs in northern and southern British English as it is a strong marker of accent. For example, F02 described a conversation with southern friends at university that centred on whether the 'right' way to pronounce words like *bath* and *grass* was with the long, back vowel [A] (SSBE) or the short vowel [a] (northern English). Trudgill (1986) also describes this difference as being one that is highly salient, so much so that "northerners would rather drop dead than produce these words like a southerner" (Trudgill, 1986: 18).

Furthermore, based on the results of Study 1, it was predicted that subjects in Study 2 who changed their spoken accent, would also change their perceptual categorizations to match what they produced. However, subjects changed their production of words like *bud* and *cud*, but changed their vowel categorization decisions for *could*. Subjects who changed their accent to sound more southern chose a central vowel for *could* that was closer to how SSBE speakers produce the vowel in words like *bud* and *cud*.

One explanation for these results is that subjects were performing the task using stored exemplars, but that they were using inaccurate exemplars that had been affected by category assimilation processes (e.g., Best, 1994; Best et al., 1988, 2001; Flege, 1995) or perceptual magnet effects (Iverson and Kuhl, 1995, 1996, 2000). Thus, the basic hypothesis that subjects were performing the task based on long-term memory representations may be correct, but these memory representations may be inaccurate. However, two aspects of the results in Study 2 are inconsistent with this account. First, subjects chose a central best exemplar for *could* after experience with SSBE, that was closer to how southerners produce the vowel in *bud* and *cud*, even though both northerners and southerners produce *could* using the same high-back vowel, [ʊ]. Second, there was no normalization for *bath*, as in Study 1. Southerners and northerners both use the vowels [a] and [ɑː], and speakers are very aware that differences in the lexical distribution of these vowels are a clear marker of accent (see earlier discussion; Trudgill, 1986, Wells, 1982b). Thus, northerners know that southerners produce this word with a long vowel [ɑː], and southerners know that northerners produce this word with a short vowel, [a]. However, subjects in both Study 1 and Study 2 chose vowels for *bath* based on their own accent rather than on their knowledge of what vowel would be expected based on the accent of the carrier sentence.

7.2.2. Studies of second language acquisition

Aspects of the present results are consistent with descriptions of how subjects perceive and produce foreign or foreign-accented speech. Cross-language research has shown that second language learners are unable to easily adapt their native language representations when learning a non-native language. Although some plasticity remains, learners typically make only small changes to production and perception. Likewise, subjects in Study 2 made only small adjustments to their production and perception after experience with SSBE speech.

Other aspects of the results are inconsistent with accounts of second language acquisition, though. Models of second language learning (see Chapter 3, Section 3.2.3; e.g., Best, 1994; Best et al., 1988, 2001; Flege, 1995) have suggested that subjects are unable to easily adapt their perceptual representations; even experienced listeners tend to assimilate non-native phonemes into their closest matching native language categories, and only create new categories where the non-native phoneme is perceived to be sufficiently different. Based on this evidence, one could imagine that subjects in Study 1 and Study 2 would assimilate non-native phonemes into their closest matching native language categories. However, subjects in both Study 1 and

Study 2 shifted their existing vowel categories when listening to speech produced in a non-native accent. For example, southerners in Study 1 chose a low-central vowel for *bud* and *cud* in Sheffield English carrier sentences instead of using their native [ʊ] category that exactly matches how northerners produce these vowels. Similarly, subjects in Study 2 chose a central vowel for *could* in SSBE carrier sentences, even though northerners and southerners produce this word using the same vowel category.

7.3. Towards an explanation

Although the pattern of results in Study 1 may seem idiosyncratic, they correspond closely with the changes that speakers tend to make to their own accent when they live in multidialectal environments (Trudgill, 1986; Wells, 1982b). Northerners who live in the south of England typically modify some aspects of their accent in order to fit in with southerners; they change their production of the vowel in *bud* and *cud* so that it becomes centralized (Trudgill, 1986; Wells, 1982b), much like the centralized vowel that northern subjects living in London chose as best exemplars for these words in SSBE-accented sentences in Study 1. Northerners also maintain some aspects of their regional identity; they retain their [a] vowel when producing *bath*, much like these northern subjects chose [a] for *bath* in both southern and northern English carrier sentences.

Study 2 also confirmed that there was a link between production and perception. Subjects chose best exemplar locations in the Goodness Optimization Task that matched how they produce these vowels. Thus, subjects who produced southern vowels in words like *bud* and *cud* chose best exemplar locations for these vowels that were closer to how SSBE speakers produce these vowels. However, the evidence for this link not was as strong as had been expected based on the results of

Study 1; the changes in production occurred in the vowels in *bud* and *cud*, but the change in perception occurred in the vowel in *could*.

Changes in production in Study 2 were also greater than changes in perception. This suggests that at least in this case, changes in production may lead changes in perception. That is, learners may find it easier to achieve native-like pronunciation in a non-native accent or language, than to achieve native-like performance on a vowel categorization task, for example. This finding is supported by evidence from studies of L2 acquisition that have indicated that production of a non-native sound is easier and/or faster to learn than perception of the same sounds (e.g., Goto, 1971; Tees and Werker, 1984; Flege and Eefting, 1984; Bohn and Flege, 1997). Bohn and Flege (1997) demonstrated that experienced learners of an L2 are able to produce a new vowel contrast like native speakers of the L2, but that perceptual abilities for a new vowel contrast may lag behind even after several years of L2 experience.

It is possible that the perception of a new vowel contrast in both a non-native accent and L2 is more resistant to the effects of experience than production because production is more subject to social control than perception. That is, perception is a private matter, one that is not shared. Production however, is a social act; speakers may want to conform to the production norms of the non-native accent or language community in order to facilitate communication, and perhaps more importantly in this study, to fit in and identify themselves with the new community. Thus, non-native listeners may be able to function adequately in perception by making only small changes to their native perceptual representations or maybe even without adapting their native perceptual processes, but may make larger changes in production in order to try to achieve native-like pronunciation (see also Bohn and

Flege, 1997). It is possible then that the subjects in Study 2 had altered their production to a greater extent, because although they were able to perceive SSBE speakers accurately by making only small changes to their native perceptual representations, they needed to make greater changes to their production in order to fit in with SSBE speakers and avoid being stigmatized as outsiders in their new community.

Such an account would not be wholly consistent with either a strong version of Motor Theory or Direct Realism, nor a General Auditory position. The results from both Study 1 and Study 2 are to a certain extent consistent with Motor Theory's and Direct Realism's claims that subjects perceive speech in terms of their own articulatory gestures. That is, subjects' acquisition of new articulatory targets for *bud* and *cud* when changing their production to better match SSBE speakers, may have directly changed their best exemplar locations in the Goodness Optimization Task. However, it is not clear how Motor Theory or Direct Realism would explain the findings in Study 2 that changes in perception lagged behind production, or more importantly, that changes in perception did not correspond to the changes in production. Equally, whilst these findings might be considered to be consistent with a General Auditory position, in that the different changes in production and perception could be taken as evidence for the two domains operating independently, it is not clear how this position would explain the finding that subjects who were judged to have a more southern accent overall also chose more southern best exemplar locations and were better at identifying SSBE speech in noise, indicating that there is a link between production and perception. It is perhaps more likely that whilst there is a link between the production and the perceptual domains, and changes in one domain can affect the other, changes in each domain might operate at different rates and be subject to different constraints and influences, such as the sociolinguistic factors described above (see also Bohn and Flege, 1997).

Such sociolinguistic factors may also play an important role in explaining why northerners in Study 1 and Study 2 changed their production and perception of some words that are produced differently in SSBE but not others. Northerners in Study 1 and Study 2 did not change their vowel categorization decisions for *bath* in SSBE carrier sentences, even though some changed their best exemplar location for words like *bud* and *cud*. Similarly, subjects in Study 2 who changed their accent did not change their production of *bath*, even though they changed their production of the vowel in words like *bud* and *cud*. It is possible that this is because *bath* is considered to be a strong marker of accent. Thus, northern subjects may choose to retain their native vowel in *bath* in order to continue to signal belonging to their native accent community (Foulkes and Docherty, 1999; Watt, 1998), and because the adoption of the SSBE [ɑː] category for these words may be seen by other northerners to be disloyal to their native accent community (Hickey, 1999). In contrast, the adoption of a new, centralized vowel for words such as *bud* and *cud* that is closer to the SSBE [ʌ] category may not be associated as strongly with SSBE by other northern English speakers, and may thus be readily available to be adopted (Hickey, 1999).

Study 2 also demonstrated that subjects who had a strong motivation to adapt changed their accent more. Subjects who had a friendship group made up predominantly of SSBE speakers and who wanted to be identified with this community changed their accent more than those whose friendship groups were made up of speakers of similar accents and who had no motivation to change their accent. This is supported by evidence from cross-language research: Second language learners, who were highly motivated and had a strong desire to sound like a native speaker, were judged to have less of a foreign accent that those that had little motivation to learn, regardless of whether they were early or late learners (Piske et al., 2001).

7.4. Final conclusions

The findings demonstrate that subjects are able to adjust both their perception and production after experience with a non-native accent. Subjects in Study 1 chose different best exemplar locations for *bud* and *cud* – but not *bath* – in SSBE and Sheffield English carrier sentences, indicating that subjects were able to normalize these vowels for accent. However, these perceptual changes were slow to emerge; subjects in Study 2 made only small adjustments to their best exemplar locations after 9 months experience of living in a multidialectal environment.

These patterns of normalization could not be readily explained by existing theories of speech perception. Exemplar theory could not easily account for the changes in subjects' best exemplar locations, and the results were also inconsistent with models of second language learning that have suggested learners assimilate non-native speech sounds into their closest matching native categories (e.g., Best 1994; Best et al., 1988, 2001; Flege, 1992, 1995). Instead, the patterns of normalization corresponded to changes in production that speakers have been reported to make as a result of sociolinguistic factors when living in a multidialectal environment (Trudgill, 1986; Wells, 1982b).

However, the evidence for a link between production and perception in Study 2 was not as strong as had been expected based on the results of Study 1. Although subjects who chose more southern best exemplars for *bud* and *cud* overall, produced versions of these vowels that were judged to sound more southern, the changes in production occurred in different vowels to the change in perception. Subjects did not change their best exemplar location for *cud* and *bud*, but instead changed their best exemplar location for *could* so that it was similar to the exemplar that they chose for *cud* and *bud*. The reason for this pattern of results is not clear. Even before experience of interacting with SSBE speakers (i.e., T1), subjects chose a relatively central best exemplar for *bud* and *cud* in the Goodness Optimization Task that was similar to the way in which they produced this vowel. Consequently, it was hypothesized that they chose the same best exemplar for *could* after experience with SSBE, because they were hypercorrecting. One shortcoming of Study 2 is that it was not possible to collect production data of subjects interacting with SSBE speakers. It is thus possible that in this multidialectal situation, subjects would also produce similar forms for words like *could* with a more central vowel, similar to the one they chose as a best exemplar in the Goodness Optimization Task. It would be valuable to test this hypothesis in further research.

The changes in production were also greater than those in perception. This finding adds to the results of studies of L2 acquisition that have suggested that the production of L2 sounds is easier and/or faster to learn than the perception of those sounds (Bohn and Flege, 1997; Flege and Eefting, 1987). However, it is unclear from this study and previous research whether or not subjects are able to change their perception to the same extent. It is thus possible that production changes faster than perception, and that after more experience with SSBE these subjects might show greater changes in perception to match those found in production. Equally, it is possible that even after more experience, subjects will show no more change in perception. This would provide further support for the explanation put forward by Bohn and Flege (1997); that perception is more resistant to experience with a non-native accent or language than production, because production is more subject to social control.

In conclusion, the results demonstrate that subjects are able to adapt to a non-native accent at a late stage in their linguistic development, although these changes are small, complex and slow to emerge. The changes in production and perception also appear to be linked to sociolinguistic factors, such as the degree to which a person wishes to identify him-/herself with a given community. In this respect, any model which is to fully account for how subjects adjust to regional variation, and possibly a second language, must account not only for the ability to adjust one's production and perception processes, but also one's motivation to fit in with that particular community.

8. Bibliography

Adank. P. (2003). *Vowel normalization: A perceptual-acoustic study of Dutch vowels.* Unpublished PhD dissertation, University of Nijmegen, The Netherlands.

Adank, P., Smits, R. & van Hout, R. (2004). "A comparison of vowel normalization procedures for language variation research." *Journal of the Acoustical Society of America* **116**: 3099-3107.

Ainsworth, W. A. and Paliwal, K. K. (1984). "Correlation between the production and perception of the English glides /w, r, l, j/." *Journal of Phonetics* **12**: 237-243.

Baker, R. J. & Rosen, S. (2001). "Evaluation of maximum-likelihood threshold estimation with tone-in-noise masking." *British Journal of Audiology* **35**: 43-52.

Bailey, P. J. and Haggard, M. P. (1973). "Perception and production: Some correlations on voicing of an initial stop." *Language and Speech* **16**: 189-195.

Bailey, P. J. and Haggard, M. P. (1980). "Perception-production relations in the voicing contrast for initial stops in 3-year-olds." *Phonetica* **37**: 377-396.

Bench, J., Kowal, A. & Bamford, J. (1979). "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children." *British Journal of Audiology* **13**: 108-112.

Best, C. T. (1994). "The Emergence of Native-Language Phonological Influences in Infants: A Perceptual Assimilation Model." In J. C. Goodman and H. C. Nusbaum (Eds.), *The Development of Speech Perception* (pp.167-244). Cambridge, Mass., MIT Press.

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

Best, C.T., Faber, A. & Levitt, A. (1996). "Assimilation of non-native vowel contrasts to the American English vowel system". *Journal of the Acoustical Society of America* 99(4): 2602-2603.

Best, C. T., McRoberts, G. W. & Goodell, E. (2001). “Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system.” *Journal of the Acoustical Society of America* **109**: 775-794.

Best, C. T., McRoberts, G. W. & Sithole, N. N. (1988). “The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants.” *Human Perception and Performance* **14**: 345-360.

Boersma, P. &Weenink, D. (2004). Praat: doing phonetics by computer (Computer software) (University of Amsterdam, The Netherlands).

Bohn, O.-S. & Flege, J. E. (1993). “Perceptual switching in Spanish/English bilinguals.” *Journal of Phonetics* **21**: 267-290.

Bohn, O.-S. & Flege, J.E. (1997). "Perception and production of a new vowel category by adult second language learners" in James, A. & Leather, J. (Eds.) *Second-language speech: Structure and process*. Mouton de Gruyter: Berlin.

Bongaerts, T., van Summeren, C., Planken, B. & Schils, E. (1997). “Age and ultimate attainment in the pronunciation of a foreign language.” *Studies in Second Language Acquisition,* **14**: 131-158.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R. & Tohkura, Y. (1997). “ Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* **101**: 2299-2310.

Chambers, J. K. (1992). "Dialect Acquisition." *Language* **68**: 678-705.

Chambers, J.K. and Trudgill, P. (1980). *Dialectology*. Cambridge University Press.

Clark, J. and Yallop, C. (1996). *An Introduction to Phonetics and Phonology, 2nd Edition. Blackwell: Oxford.*

Clarke, C. M. (2002). *Perceptual Adjustment to Foreign-Accented English With Short Term Exposure*. Paper presented at the 7th International Conference on Spoken Language Processing (ICSLP), Denver, CO, U.S.A.

Clarke, C. M. (2004). "Listening to a non-native speaker: Adaptation and generalization." Paper presented at the 147th Meeting of the Acoustical Society of America, New York, NY.

DeKeyser, R. (2002). "The robustness of critical period effects in second language acquisition." *Studies in Second Language Acquisition* **22**: 499-534.

Diehl, R. L., Lotto, A.J. & Holt, L. L. (2004). "Speech Perception." *Annual Review of Psychology* **55**: 149-179.

Docherty, G. J. & Foulkes, P. (1999). "Derby and Newcastle: Instrumental phonetics and variationist studies." In P. Foulkes and G. J. Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles* (pp.47-71). London: Arnold Press.

Eisner, F. & McQueen, J. (in press). "The Specificity of Perceptual Learning in Speech Processing." *Perception and Psychophysics.*

Elman, J. L., Diehl, R. L. & Buchwald, S. E. (1977). "Perceptual switching in bilinguals." *Journal of the Acoustical Society of America* **62**: 971-974.

Evans, B. G. (2001). *An investigation into the identification of four varieties of English.* Unpublished master's thesis. University of Newcastle upon Tyne, Newcastle upon Tyne, U.K..

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**: 352-361.

Fabricius, A. (2002). "Ongoing change in modern RP: Evidence for the disappearing stigma of t-glotalling". *English World-Wide* **23**: 115-136.

Fadiga L., Fogassi L., Gallese V., and Rizzolatti G. (2000). "Visuomotor neurons: ambiguity of the discharge or 'motor' perception?" *International Journal of Psychophysiology* **35**: 165-177.

Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). "Speech listening specifically modulates the excitability of tongues muscles: a TMS study." *European Journal of Neuroscience* **15**: 399-402.

Flege, J. E. (1991). "Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language." *Journal of the Acoustical Society of America* **89**: 395-411.

Flege, J. E. (1992). "The intelligibility of English vowels spoken by British and Dutch talkers." In R. D. Kent (Ed.), *Intelligibility in Speech Disorders: Theory, Measurement and Management* (pp.157-232). John Benjamins: Amsterdam/Philadelphia.

Flege, J. E. (1995). "Second language speech learning: Theory, findings, and problems." In W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp.233-277). York Press: Baltimore.

Flege, J. E. (2002). “Interactions between the Native and Second-language Phonetic Systems.” In P. Burmeister, T. Piske, & A. Rohde (Eds.), *An integrated View of Language Development: Papers in Honour of Henning Wode*. Trier: Wissenschaftlicher Verlag Trier.

Flege, J. E. (2003). “Assessing constraints on second-language segmental production and perception.” In A. Meyer & N. Schiller (Eds.) *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*. Berlin: Mouton de Gruyter.

Flege, J. E. (2004). “Age of learning and second language speech”. In D. Birdsong (Ed.) *Second Language Acquisition and the Critical Period Hypothesis*. Lawrence Erlbaum Associates: London.

Flege, J. E. & Eefting, W. (1987). “Cross-language switching in stop consonant perception and production by Dutch speakers of English.” *Speech Communication* **6**: 185-202.

Flege, J. E., Frieda, E. & Nozawa, T. (1997). “Amount of native-language (L1) use affects the pronunciation of an L2.” *Journal of Phonetics* **25**: 169-186.

Flege, J.E. & Hillenbrand, J. (1984). "Limits on phonetic accuracy in foreign language speech production." *Journal of the Acoustical Society of America* 76(3): 708-722.

Flege, J. E. & MacKay, I. R. A. (2004). “Perceiving vowels in a second language.” *Studies of Second Language Acquisition* **24**: 1-34.

Flege, J. E., Mackay, I. R. A. & Meador, D. (1999). “Native Italian speakers' perception and production of English vowels.” *Journal of the Acoustical Society of America* **106**: 2973-2987.

Flege, J. E., Schirru, C. & MacKay, I. R. A. (2002). "Interaction between the native and second language phonetic subsystems." *Speech Communication* **40**: 467-491.

Foulkes, P. & Docherty, G.J. (1999). "Urban Voices – Overview." In P. Foulkes & G. J. Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles*. London: Arnold Press.

Fowler, C. A. (1981). "Production and perception of coarticulation among stressed and unstressed vowels." *Journal of Speech, Hearing and Language Research* **24**: 127-139.

Fowler, C. A. (1986). "An event approach to the study of speech perception from a direct-realist perspective." *Journal of Phonetics* **14**: 3-28.

Fujisaki, H. & Kawashima, T. (1968). "The roles of pitch and higher formants in the perception of vowels." *IEEE Transactions on Audio and Electroacoustics AU-16*: 73-77.

Gerstmann, L. (1968). "Classification of self-normalized vowels." *IEEE Transactions on Audio and Electroacoustics AU-16*: 78-80.

Goldinger, S. D. (1996). "Words and Voices: Episodic Traces in Spoken Word Identification and Recognition Memory." *Experimental Psychology* **22**: 1166-1183.

Goldinger, S. D. (1998). "Echoes of Echoes? An Episodic Theory of Lexical Access." *Psychological Review* **105**: 251-279.

Goldinger, S. D., Kleider, H. M. & Shelley, E. (1999). " The marriage of perception and memory: Creating two-way illusions with words and voices." *Journal of Memory and Cognition* **27**: 328-338.

Goto, H. (1971). "Auditory perception by normal Japanese adults of the sounds "L" and "R". *Neuropyschologia* 9: 317-323.

Halle, M. (1985). "Speculations about the representations in memory." In V. A. Fromkin (Ed.), *Phonetic linguistics*. New York: Academic Press.

Hawkins, S. (2003). "Contribution of Fine Phonetic Detail to Speech Understanding". Paper presented at International Congress of Phonetic Sciences (ICPhS), Barcelona, Spain.

Hazan, V. L. & Boulakia, G. (1992). "Perception and Production of a voicing contrast by French-English bilinguals." *Language and Speech* **36:**17-38.

Hickey, R. (1999). "Dublin English: current changes and their motivation". In P. Foulkes and G. J. Docherty (Eds.) *Urban Voices: Accent Studies in the British Isles.* London: Arnold Press.

Hillenbrand, J., Getty, L. A., Clark, M. J. & Wheeler, K. (1995). "Acoustic characteristics of American English vowels." *Journal of the Acoustical Society of America* **97**: 3099-3111.

Huckvale, M. (2003). Speech Filing System (computer software) (University College, London. U.K.).

Hughes, A., Trudgill, P. & Watt, D. (2005). *English Accent and Dialects: An Introduction to Social and Regional Varieties of English in the British Isles,* 4th edition. Hodder Arnold.

Iverson, P. & Evans, B. G. (2003). " A Goodness Optimization Method for Investigating Phonetic Categorization." Paper presented at the International Congress of Phonetic Sciences (ICPhS), Barcelona, Spain.

Iverson, P. & Kuhl, P. K. (1995). "Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling." *Journal of the Acoustical Society of America* **97**: 553-562.

Iverson, P. & Kuhl., P. K. (1996). "Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/." *Journal of the Acoustical Society of America* **99**: 1130-1140.

Iverson, P. & Kuhl, P. K. (2000). "Perceptual magnet and phoneme boundary effects in speech perception: do they arise from a common mechanism?" *Perception and Psychophysics* **62**: 874-883.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A. & Siebert, C. (2004). "A perceptual interference account of acquisition difficulties for non-native phonemes." *Cognition* **87**: B47-B57.

Johnson, J. & Newport, E. (1989). "Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language." *Cognitive Psychology* **21**: 60-99.

Johnson, K. (1997). "Speech Perception without Speaker Normalization: An Exemplar Model." In K. Johnson & J. W. Mullenix (Eds.) *Talker Variability in Speech Processing* (pp.145-165). Academic Press: San Diego.

Johnson, K., Flemming, E. & Wright, K. (1993). "The hyperspace effect: Phonetic targets are hyperarticulated." *Language* **69**: 505-28.

Kerswill, P. & Williams, A. (2000). "Creating a new town koine: Children and language change in Milton Keynes." *Language in Society* **29**: 65-115.

Klatt, D. H. (1989). "Review of Selected Models of Speech Perception." In W. Marslen-Wilson (Ed.), *Lexical Representation and Process.* MIT Press" Cambridge, Mass..

Klatt, D. H. & Klatt, L. C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers." *Journal of the Acoustical Society of America* **87**: 820-857.

Kuhl, P. K. (1979). "Speech perception in early infancy: Perceptual constancy for spectrally dissimilar vowel categories." *Journal of the Acoustical Society of America* **66**: 1668-1679.

Kuhl, P.K. (1981). "Discrimination of speech by nonhuman animals: basic auditory sensitivities conducive to the perception of speech-sound categories." *Journal of the Acoustical Society of America* **95**: 340-349.

Kuhl, P. K. (1983). "Perception of Auditory Equivalence Classes for Speech in Early Infancy." *Infant Behavior and Development* **6**: 263-285.

Kuhl, P. K. (2000). "A new view of language acquisition." *Proceedings of the National Academy of Sciences, U.S.A.* **97**: 11850-11857.

Kuhl, P. K., Tsao, F-M. & Liu, H-M. (2003). "Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning." *Proceedings of the National Academy of Sciences, U.S.A.* **100**: 9096-9101.

Kuhl, P. K. & Meltzoff, A.M. (1996). "Infant vocalizations in response to speech: Vocal imitation and developmental change." *Journal of the Acoustical Society of America* **100**: 2425-2438.

Kuhl, P. K. & Miller, J. D. (1975). "Speech perception by the chinchilla: Voiced-voiceless distinction in alveolar plosive consonants." *Science* **190**: 69-72.

Kuhl, P. K. & Miller, J. D. (1978). "Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *Journal of the Acoustical Society of America* **63**: 905-917.

Kuhl, P. K., Williams, K.A., Lacerda, F., Stevens, K. & Lindblom, B. (1992). "Linguistic Experience Alters Phonetic Perception in Infants by 6 Months of Age." *Science* **255**: 606-608.

Labov, W. (1972). *Language in the Inner City: Studies in the Black English Vernacular*. Philadephia: University of Pennsylvania Press.

Labov, W. and Ash, S. (1997). "Understanding Birmingham." In C. Bernstein, T. Nunally, and R. Sabino (Eds.) *Language Variety in the South Revisited.* Alabama: Tuscaloosa.

Ladefoged, P. & Broadbent, D. (1957). "Information conveyed by vowels." *Journal of the Acoustical Society of America* **29**: 98-104.

Lenneberg E., (1967). *Biological Foundations of Language*. New York: John Wiley & Sons, Inc..

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics." *Journal of the Acoustical Society of America* **49**: 467-477.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967). "Perception of the speech code." *Psychological Review* **74**: 431-461.

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

Lobanov, B. M. (1971). "Classification of Russian vowels spoken by different speakers." *Journal of the Acoustical Society of America* **49**: 606-608.

Logan, J. S., Lively, S. E. & Pisoni, D. B. (1991). "Training Japanese listeners to identify /r/ and /l/: A first report." *Journal of the Acoustical Society of America* **89**: 874-886.

Marean, G. C., Werner, L. A. & Kuhl, P. K. (1992). "Vowel categorization by very young infants." *Developmental Psychology* **28**: 396-405.

Meador, D., Flege, J. & MacKay, I178 (2000). "Factors affecting the recognition of words in a second language." *Bilingualism: Language and Cognition* **3**" 55-67.

Miller, R. L. (1953). "Auditory tests with synthetic vowels." *Journal of the Acoustical Society of America* **25**: 114-121.

Milroy, L. (1987a). *Observing and Analysing Natural Language*. Oxford: Blackwell.

Milroy, L. (1987b). *Language and Social Networks* (2nd Ed.). Oxford: Blackwell.

Moore, B. C. J., Glasberg, B. R. & Baer, T. (1997)."A Model for the Prediction of Thresholds, Loudness, and Partial Loudness." *Journal of the Audio Engineering Society* **45**: 224-240.

Nearey, T. M. (1989). "Static, dynamic, and relational properties in vowel perception." *Journal of the Acoustical Society of America* **85**: 2088-2113.

Newbrook, M. (1999). "West Wirral: norms, self-reports and usage." In P. Foulkes and G. J. Docherty (Eds.) *Urban Voices: Accent Studies in the British Isles*. London: Arnold.

Newman, R. S. (2003). "Using links between speech perception and speech production to evaluate different acoustic metrics: A preliminary report". *Journal of the Acoustical Society of America* **113**: 2850-2860.

Nygaard, L. C. & Pisoni, D. B. (1998). "Talker-specific Learning in Speech Perception." *Perception and Psychophysics* **60**: 355-376.

Nygaard, L. C., Sommers, M. S. & Pisoni, D. B. (1995). "Effects of stimulus variability on perception and representation of spoken words in memory." *Perception and Psychophysics* **57**: 989-1001.

Palmeri, T. J., Goldinger, S. D. & Pisoni, D. B. (1993). "Episodic encoding of voice attributes and recognition memory for spoken words." *Journal of Experimental Psychology: Learning, Memory, and Cognition* **19**: 309-328.

Patkowski, M. (1998). "Age and accent in a second language: A reply to James Emil Flege." *Applied Linguistics* **11**: 73-89.

Perkell, J. S., Guenther, F., Lane, H., Matthies, M. L., Stockmann, E. Tiede, M. and Zandipour, M. (2004). "The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts." *Journal of the Acoustical Society of America* **116**: 2338-2344.

Pierrehumbert, J. B. (2001). "Exemplar dynamics: Word frequency, lenition and contrast." In J. Bybee & P. Hopper (Eds.), *Frequency effects and emergent grammar*. Amsterdam: John Benjamins.

Piske, T., Mackay, I. R. A. & Flege, J. E. (2001). "Factors affecting degree of foreign accent in an L2: a review." *Phonetics* **29**: 191-215.

Pisoni, D. B. (1977). "Identification and discrimination of the relative onset time of two component tones: Implications for voicing perception in stops." *Journal of the Acoustical Society of America* **62**: 1352-1361.

Pisoni, D. B. (1997). "Some Thoughts on 'Normalization' in Speech Perception." In K. Johnson & J. W. Mullenix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). Academic Press: San Diego.

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. H. (1992). *Numerical recipes in C: The art of scientific computing*. Cambridge; New York: Cambridge University Press, 2nd Ed..

Przedlacka, J. (2002). *Estuary English? A Sociophonetic Study of Teenage Speech in the Home Counties.* Oxford: Peter Lang.

Rizzolatti, G. & Arbib, M. A. (1998). "Language within our grasp." *Trends in Neuroscience* **21**: 188-194.

Sangster, C. (2002). *Inter- and Intra-speaker variation in Liverpool English: A sociophonetic study.* Unpublished PhD dissertation, University of Oxford, Oxford, U.K.

Sankoff, G. (2004). "Adolescents, young adults and the critical period: Two case studies from 'Seven Up'." In C. Fought (Ed.) *Sociolinguistic Variation*. Oxford: Oxford University Press.

Scott, D. R. & Cutler, A. (1984). "Segmental Phonology and the Perception of Syntactic Structure". *Journal of Verbal Learning and Verbal Behavior* **23**: 450-466.

Scovel, T. (1988). *A Time to Speak: A Psycholinguistic Inquiry into the Critical Period for Human Speech.* Cambridge, MA: Newbury House Publishers.

Sebastian-Galles, N. & Soto-Faraco, S. (1999). "On-line processing of native and non-native phonemic contrasts in early bilinguals." *Cognition* **72**: 111-123.

Shankweiler, D. P., Strange, W., & Verbrugge, R. R. (1976). "Speech and the problem of perceptual constancy." In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, knowing: Toward an ecological psychology*. Hillsdale, NJ: Erlbaum.

Smith, R. (2003). "Influence of Talker-Specific Phonetic Detail on Word Segmentation." Paper presented at the International Congress of Phonetic Sciences (ICPhS), Barcelona, Spain.

Stevens, K. N. & Klatt, D. H. (1974). "Role of formant transitions in the voiced-voiceless distinction for stops." *Journal of the Acoustical Society of America* **55**: 653-659.

Stoddart, J., Upton, C. & Widdowson, J.D.A. (1999). "Sheffield dialect in the 1990s: revisiting the concept of NORMs". In Foulkes, P. & Docherty, G. (Eds.) *Urban Voices: Accent Studies in the British Isles* (pp. 163-184). London: Arnold.

Strange, W. & Dittmann. S (1984). "Effects of discrimination traning on the perception of /r-l/ by Japanese adults learning English. *Perception and Psychophysics* 36: 131-145.

Stuart-Smith, J. (1999). "Glasgow: accent and voice quality". In P. Foulkes and G. J. Docherty (Eds.) *Urban Voices: Accent Studies in the British Isles*. London: Arnold.

Tees, R.C. & Werker, J.F. (1984). "Perceptual flexibility: Maintenance or recovery of the ability to discriminate nonnative speech sounds". *Canadian Journal of Psychology* 38: 579-590.

Tollfree, L. (1999). "South East London English: discrete versus continuous modelling of consonantal reduction". In Foulkes, P. & Docherty, G. (Eds.) *Urban Voices: Accent Studies in the British Isles* (pp. 163-184). London: Arnold.

Torgersen, E. (1997). *Some phonological innovations in south-eastern British English*. Unpublished master's thesis. University of Bergen, Bergen, Norway.

Trudgill, P. (1986). *Dialects in Contact*. Oxford, Basil Blackwell.

Upton, C. and Widdowson, J. D. A. (1996). *An Atlas of English Dialects*. Oxford, Oxford University Press.

Verbrugge, R. R., Strange, W., Shankweiler, D. P. & Edman, T. R. (1976). "What information enables a listener to map a talker's vowel space?" *Journal of the Acoustical Society of America* **60**: 198-212.

Vihman, M. M. (1991). "Ontogeny of phonetic gestures." In I.G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception*. Hillsadale NJ: Lawrence Erlbaum.

Vihman, M. M. & Nakai, S. (2003). "Experimental Evidence for an Effect of Vocal Experience on Infant Speech Perception." Paper presented at the 15th International Congress of Phonetic Sciences (ICPhS), Barcelona, Spain.

Watt, D. J. L. (1998). *Variation in Change in the Vowel System of Tyneside English.* Unpublished PhD dissertation, University of Newcastle, U.K..

Wells, J. C. (1982a). *Accents of English: An Introduction,* (Vol. 1). Cambridge, Cambridge University Press.

Wells, J. C. (1982b). *Accents of English: The British Isles,* (Vol. 2). Cambridge, Cambridge University Press.

Wells, J. C. (1984). "The cockneyfication of RP?" In Melchers, Gunnel and Johannesson, Nils-Lennart (Eds.) *Non-standard Varieties of Language.* (pp. 198-205). Stockholm: Almqvist & Wiksell.

Williams, A. & Kerswill, P. (1999). "Dialect levelling: change and continuity in Milton Keynes, Reading and Hull". In P. Foulkes and G. J. Docherty (Eds.), *Urban Voices: Accent Studies in the British Isles* (pp.141-162). London: Arnold Press.

White, L. & Genessee, F. (1996). "How native is near-native? The issue of ultimate attainment in adult second language acquisition." *Second Language Research* **12**: 238-265.

9. Appendices

9.1. Appendix A: Reading passage, Study 1.

By twelve o'clock next day, Harry's trunk was packed with his school things, and all his most prized possessions – the Invisibility Cloak he had inherited from his father, the broomstick he had got from Sirius, the enchanted map of Hogwarts he had been given by Fred and George Weasley last year. He had emptied his hiding place under the loose floorboard of all food, double-checked every nook and cranny of his bedroom for forgotten spellbooks or quills, and taken down the chart on the wall counting the days down to September the first, on which he liked to cross off the days remaining until his return to Hogwarts.

The atmosphere inside number four Privet Drive was extremely tense. The imminent arrival at their house of an assortment of wizards was making the Dursleys uptight and irritable. Uncle Vernon had looked downright alarmed when Harry informed him that the Weasleys would be arriving at five o'clock the very next day.

'I hope you told them to dress properly, these people,' he snarled at once. 'I've seen the sort of stuff your lot wear. They'd better have the decency to put on normal clothes, that's all.'

Harry felt a slight sense of foreboding. He had rarely seen Mr or Mrs Weasley wearing anything that the Dursleys would call 'normal'. Their children might don Muggle clothing during the holidays, but Mr and Mrs Weasley usually wore long robes in varying states of shabbiness. Harry wasn't bothered about what the neighbours would think, but he was anxious about how rude the Dursleys might be to the Weasleys if they turned up looking like their worst idea of wizards.

From J.K. Rowling (2000), *Harry Potter and the Goblet of Fire*. Bloomsbury.

Appendix B: Subject Questionnaire, Study 1

University College London

Department of Phonetics and Linguistics
Wolfson House, 4 Stephenson Way, London NW1 2HE

Miss Bronwen G. Evans

10th January 2002, Version A

Questionnaire

The information which you give in this questionnaire will be used to help interpret the results of the study. It will not be used for any other purpose and will remain confidential.

Name: .... **Date of Birth**.....................

Occupation: .... **Sex:** M/F

Current Home Address: ....

Have you always lived at your current address? Yes / No

If not, please give details of the other places in which you have lived, including the length of time you spent there and your age at the time of leaving. (Include places of study e.g. university, boarding school etc. as well as places of residence).

....

....

....

....

....

Do you speak any foreign languages? Yes / No

If yes, please give details of the language(s) spoken and your level of ability.

....

....

....

....

....

....

....

9.2. Appendix C: Subject Questionnaire, Study 2.

University College London

Department of Phonetics and Linguistics

Wolfson House, 4 Stephenson Way, London NW1 2HE

Miss Bronwen G. Evans

10th January 2002, Version A

Questionnaire

The information which you give in this questionnaire will be used to help interpret the results of the study. It will not be used for any other purpose and will remain confidential.

Name: .... **Date of Birth**.....................

University attended: .... **Sex:** M/F

Current Home Address: ....

Have you always lived at your current address? Yes / No

If not, please give details of the other places in which you have lived, including the length of time you spent there and your age at the time of leaving. (Include places of study e.g. university, boarding school etc. as well as places of residence).

....

....

....

....

....

Do you speak any foreign languages? Yes / No

If yes, please give details of the language(s) spoken and your level of ability.

....

....

....

....

....

....

....

9.3. Appendix D: Arthur the Rat

Once there was a young rat named Arthur, who could never make up his mind. Whenever his friends asked him if he would like to go out with them, he would only answer, “I don't know.” He wouldn't say “yes” or “no” either. He would always shirk at making a choice.

Even his aunt Helen said to him, “Now look here. No one is ever going to care for you if you carry on like this. You have no more mind than a blade of grass.”

One rainy day, the rats heard a great noise in the loft. The pine rafters were all rotten, so that the barn was rather unsafe. At last the joists gave way and fell to the ground. The walls shook and all the rats' hair stood on end with fear and horror. “This won't do,” said their elderly captain. “I'll send out scouts to search for a new home.”

Within five hours the ten scouts came back and said, “We found a stone house where there is room and board for us all. There is a kindly horse named Nelly, a cow, a calf, and a garden with an elm tree.” The rats crawled out of their little houses and stood on the floor in a long line, ready to march away.

Just then the old captain saw Arthur. “Stop,” he ordered the others coarsely. “You are coming, of course?” “I'm not certain,” said Arthur, undaunted. “The roof may not come down yet.” “Well,” said the angry old rat, “we can't wait for you to join us. Right about face. March!”

Arthur stood and watched them hurry away. “I think I'll go tomorrow,” he calmly said to himself, but then again “I don't know; it's so nice and snug here.”

That night there was a big crash. In the morning some men — with some boys and girls — rode up and looked at the barn. One of them moved a board and he saw a young rat, quite dead, half in and half out of his hole. Thus the shirker got his due.

9.4. Appendix E: A study of young people from Tyneside

Young people from Tyneside are reluctant to cast themselves as old-fashioned, as old-fashionedness in an urban industrial setting tends to be congruent with poverty. The occupations, lifestyles and pastimes that had traditionally been associated with Tynesiders in the past are unlikely to hold many attractions for Tynesiders of any social class today. However, complete subscription to middle-class values among young working-class people is unlikely, since loyalty to the local community will continue to be a potent force in terms of the shaping of social behaviour. [...]

Many young Tyneside people, it appears, wish to throw off some of the baggage of the perceived parochiality, uncouthness and uneducatedness that has been visited upon them by outsiders for the last two hundred years, and to present themselves as part of a wider, more worldly, more dynamic urban British population. On the other hand, while they may resent the caricature of the 'canny Geordie' which has entered the British mindset as a recognisable stereotype, they are unlikely to forgo the well-deserved reputation for warmth, friendliness, hospitality, straightforwardness, honesty and unpretentiousness on which they pride themselves.

From D. J. L. Watt (1998), *Variation in Change in the Vowel System of Tyneside English.* Unpublished PhD dissertation, University of Newcastle, U.K..