Approaching intonational distance and change

Jennifer Niamh Sullivan, B.A., M.Litt.

**A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy**

**to
School of Philosophy, Psychology and Language Sciences
College of Humanities and Social Sciences
University of Edinburgh**

September 2010

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Jennifer Niamh Sullivan

Abstract

The main aim of this thesis is to begin to extend phonetic distance measurements to the domain of intonation. Existing studies of segmental phonetic distance have strong associations with historical linguistic questions. I begin with this context and demonstrate problems with the use of feature systems in these segmental measures. Then I attempt to draw strands from the disparate fields of quantitative historical linguistics and intonation together. The intonation of Belfast and Glasgow English provides a central case study for this. Previous work suggests that both varieties display nuclear rises on statements, yet they have never been formally compared.

This thesis presents two main hypotheses on the source of these statement rises: the Alignment hypothesis and the Transfer hypothesis. The Alignment hypothesis posits that statement rises were originally more typical statement falls but have changed into rises over time through gradual phonetic change to the location of the pitch peak. The Transfer hypothesis considers that statement rises have come about through pragmatic transfer of rises onto a statement context, either from question rises or continuation rises. I evaluate these hypotheses using the primary parameters of alignment and scaling as phonetic distance measurements. The main data set consists of data from 3 Belfast English and 3 Glasgow English speakers in a Sentence reading task and Map task.

The results crucially indicate that the origin of the statement rises in Belfast and Glasgow English respectively may be different. The Glasgow statement nuclear tones show support for the Alignment hypothesis, while the Belfast nuclear tones fit best with the Transfer hypothesis. The fundamental differences between Glasgow and Belfast are the earlier alignment of the peak (H) in Glasgow and the presence of a final low (L) tonal target in Glasgow and a final high (H) target in

Belfast. The scaling of the final H in Belfast statements suggests that the transfer may be from continuation rather than from question rises.

I then present a proposal for an overall measure of intonational distance, showing problems with parameter weighting, comparing like with like, and distinguishing between chance resemblance and genuine historical connections. The thesis concludes with an assessment of the benefits that intonational analysis could bring to improving segmental phonetic distance measures.

Acknowledgements

I have been very fortunate to have been allowed to use the PPLS/Informatics high quality recording facility for Experiment II (chapter 4) and I am very grateful to Barry Campbell, sound technician, for his assistance with the recordings. I also thank the other technical staff in Linguistics and English language. There would be no experimental research without the experimental participants, who gave their time and enthusiasm to my tasks. I am immensely thankful to them.

Experiment I (chapter 4) was made possible by the availability of the IViE (Intonational Variation in English) Corpus of English varieties. I am deeply grateful to the IViE project researchers for making their data freely available for research purposes. I include official acknowledgements to them here.

Conditions of use: The IViE corpus and the associated documentation are copyright information. The speech data and the documentation cannot be copied or distributed in any format unless this paragraph is included. The speech data are available to any interested user, but only for non-commercial use. The UK ESRC and the University of Oxford make no warranty whatsoever in relation to the materials including as to accuracy, quality or fitness for any particular purpose, and accept no liability in relation to the use of the materials or anything associated with such use. Copyright ©2007 J. Coleman, E. Grabe and G. Kochanski.

I am grateful to Pauline Welby for making her Praat scripts available.

Florence Bonacina, Ute Medicus and a friend of hers kindly provided translation assistance in the preparation of a journal article during the preparation of this thesis.

For administrative assistance, I thank the PPLS administrative staff, in particular, Devon Taylor and Katie Keltie.

The University of Edinburgh has provided several valuable courses in general PhD training from which I have benefitted greatly.

Above all, my parents and family have been absolutely unwavering in their support throughout the last few years in ways which have been deeply humbling.

I would like to acknowledge John Benjamins Publishing Company[1] for granting

---

[1]Ms Patricia Leplae, Rights and Royalties Department, John Benjamins Publishing Co., P.O. Box 36224, NL-1020 ME Amsterdam, The Netherlands, www.benjamins.com.

me permission to include a copy of Sullivan and McMahon (2010) in the hard-bound version of this thesis.

My examiners, Professor Francis Nolan and Dr Warren Maguire, made several detailed recommendations for the improvement of this thesis, which I have followed gratefully. Naturally, all errors remain my own.

Contents

List of Tables

List of Figures

CHAPTER 1

What are phonetic distances?

> Speaking English means doing two things at once: saying the words, and saying the melody (Silverman and Pierrehumbert, 1990, p.72).

1.1 Introduction

If I told you that Belfast and Glasgow English once may have sounded more similar to some other varieties of British/Irish English than they do today, would you believe me? If so, would you think that we might be able to find some relics of that former similarity today?

An appropriately sceptical response might be:

- In exactly which respects did they sound similar?
- What exactly do I mean by 'similar' anyway?
- Where is the evidence?

Words versus 'melody': this is the fundamental distinction with which this thesis begins. Is it possible to measure how similar two different pronunciations of the same word are in two different English varieties (e.g. Southern British vs. U.S. pronunciations of 'tomato')? The crucial question for me though is that even if it *is* possible to do that, is it possible to measure distance in the 'melody' (intonation) between two different English varieties? Further, does distance between words or 'melody' indicate anything about historical change in word pronunciation or in intonation? What this thesis hopes to encapsulate is the relationship between sound distance and sound change. Its departure from previous work is

that it starts moving away from distance in word pronunciations to distance in the 'melody'.

One very distinctive characteristic of Belfast and Glasgow English has provided the motivation for this change in approach. This characteristic, rising intonation on statements, requires an account of 'melodic' distance. Its presence also raises questions of historical change to understand why Belfast and Glasgow English are different in this way from many other English varieties.

1.2   What is phonetic distance?

My research is concerned with the concept of distance, something that is intuitively easy but hard to quantify. My field of study is Phonetics, the branch of linguistics dealing with speech sounds (as opposed to sentence structure or meaning, for example). Following the opening distinction between words and 'melody', I make the corresponding broad divide in phonetics between the study of vowels and consonants (termed 'segments') and the study of phenomena such as intonation which operate over larger units ('suprasegmentals').

There are different levels of distance: One is the level of individual vowel and consonant differences. Another level has to do with impressions of distance between varieties, which may be based on distances between individual segments and/or intonation.[1] For example, a very salient distinction among English varieties is between those varieties in which speakers pronounce /r/ after vowels and those in which they do not (e.g. RP English [dɔˑtə] vs Std Scottish English [dɔtʰəɹ] (McMahon et al., 2005-07)).

Classifications of consonants and vowels based on their articulation in the vocal tract and on their acoustic properties have been long-established (Jakobson et al., 1952; Fant, 1960; Delattre, 1965; Ladefoged, 1975; Maddieson, 1984; Laver, 1994; IPA, 1999; Fant, 2004; Ladefoged, 2006). Particularly prominent has been the description of individual segments by sets of feature values. For example,

$$[u] = \begin{bmatrix} +high \\ +back \\ +round \end{bmatrix}$$

---

[1]Throughout this thesis I use the term 'variety' as it more neutral than 'dialect' or 'accent'.

$$[i] = \begin{bmatrix} +high \\ -back \\ -round \end{bmatrix}.$$

Despite this, researchers in phonetics have not, until recently, tended to use these classifications to quantify the relationships between segments. Within the last 10 to 15 years, this elusive measuring aspect of segmental distance has attracted the attention of researchers, who have duly come up with sometimes diverse and conflicting ways of tackling it. What is common to many of the approaches though is that distance is measured between vowels and consonants in isolated words between different languages/varieties. The relationships between vowels and consonants can be expressed by feature values, and researchers have attached numerical values to these (e.g. the comparison between [u] and [i] might score 2, as these vowels differ in their value for [back] and [round]). After deriving scores like this between pairs of individual segments, researchers have used various algorithms to get a score over pairs of words and then a further score over a whole set of word pairs between two different languages/varieties (see chapter 2 for full details).

Why should phonetic distance measurements be confined to the study of segments in the context of isolated words, though?

1.2.1 'Similarity' vs. 'Distance'

I observe that quantification in phonetics (and elsewhere) has used two terms, 'similarity' and 'distance'. Often these terms are quite interchangeable, except for representing opposite ends of the same scale. For consistency, I have chosen to use the term 'distance' throughout this thesis, except where the term 'similarity' is explicitly necessary to make a contrast. I am not in a position in this thesis to discuss the issue of whether 'similarity' and 'distance' are really just on the opposite ends of the same scale or not (see Kroeber and Chrétien, 1937; Frisch, 1996; Hayward, 2000; Kondrak, 2005).

1.2.2 Why measure phonetic distance?

The most prominent goals of existing approaches to measuring phonetic distances reveal why intonation has hardly been tackled until now. These are:

- Studying historical connections between languages/varieties, and

- Dialectolometry.

The word as a unit has been a special tool in both domains. Lexicostatistics was a prominent technique, which used counts of shared cognates[2] between pairs of languages/varieties to measure the closeness of related languages/varieties (Swadesh, 1972). It was paralleled in dialectometry by Séguy's similar technique (see ch.1 of Heeringa, 2004 and Séguy (1973)) to get a score of distance between synchronic varieties. These techniques are measurements of *Lexical (word)* distance. Measurements of *phonetic distance* have developed largely as extensions of these techniques. Given that varieties of a single language may be cognate in the majority of their vocabulary, phonetic distance measures in dialectometry provide a way of accessing the subtle differences in the pronunciation of cognate words.

Those interested in tackling historical questions of the existence and degree of relationships between languages/varieties will use whatever data and methods are most probative for this purpose. The existence of written records is immensely valuable where there is confidence that they reasonably faithfully represent pronunciations. However, written records contain nothing about intonation patterns. Thus, there is no such valued tradition of studying sound change here or even a clear method as to how to begin quantifying, and this is a real problem. This explains why intonation has been of little interest or value to these researchers. Similarly, the wide availability of recording equipment means that it is easy to record speakers from different varieties and analyse their speech today. The intonational information will of course be preserved here but suitable equipment for recording speech has only been around for less than 150 years (Morton, 2006). Arguably, dialectometrists have much more at their disposal to quantify than mere word pronunciations, but I hope this general background gives a plausible explanation as to why many have not.

This does not mean that the study of intonational distance and change in itself has nothing to learn from the traditions of sound change study and existing measuring techniques in segments. Furthermore, the challenge of intonational distance and change may allow new ways of thinking about these issues in segments. This thesis aims to show that intonation deserves a strong place in the context of phonetic distance and potential language change.

---

[2]Words/morphemes in different languages that have derived from the same source in an ancestral language e.g. Latin 'pater', English 'father'.

1.2.3 How the field has been developing

I have identified the two main fields in which measures of phonetic distance have been developed: Historical linguistics and Dialectometry. I have also reinforced the point that researchers in these fields have concentrated on segmental distances in the context of individual words. They have also favoured treating vowels and consonants as discrete segments comprised of discrete feature values. Several phonetic distance measures (and versions of these measures) now exist. However, there has been little comparison of these measures against each other, with the goal of examining how these measures are actually working. In chapter 2, I introduce a small-scale study to begin making inroads into this.

One of the most notable developments in Historical linguistics has been the rapid rise in the use of techniques borrowed from Evolutionary biology in combination with measurements of distance between languages and varieties. These techniques, broadly known as Phylogenetic techniques, were originally designed to study ancestral connections between biological taxa, though phylogenetic network techniques also display non-ancestral connections. Both trees and networks have been used with phonetic distance measurements. The networks have been increasingly used with such measurements among closely-related varieties as a more rigorous method of studying the complexity of connections between them.

1.3 Unexplored Intonation

The study of intonation lacks an established framework for measurements and is not preserved in writing. Intonation is also less obviously broken down into discrete segment-like pieces or features. Thus it is less clear which elements should be compared with which. These are the primary reasons why intonation has been almost untouched in the context of measurements between varieties. By contrast, current phonetic analysis on intonation is focused on extremely fine gradient acoustic measurements. Can this incongruency with discrete segmental measurements be reconciled?

1.3.1 Introduction to intonation

Speakers do not speak robotically, uttering words on a monotone. When I refer to intonation in this thesis, I am referring primarily to the speaker's use of pitch

to express meaning at the *utterance* level (Ladd, 2008). I wish to distinguish this from the use of pitch to express meaning at the *word* level (lexical tone, lexical accent).[3] A simple example of utterance level meaning might be the distinction between a statement and a question. Figure 1.1 below gives an example of the statement 'It's a war memorial' with a pitch fall on 'war memorial'. Figure 1.2 gives an example of the question 'Did she see the war memorial'? with the pitch rising on 'war memorial'. The meaning of the individual words is exactly the same, despite the intonational contrast.
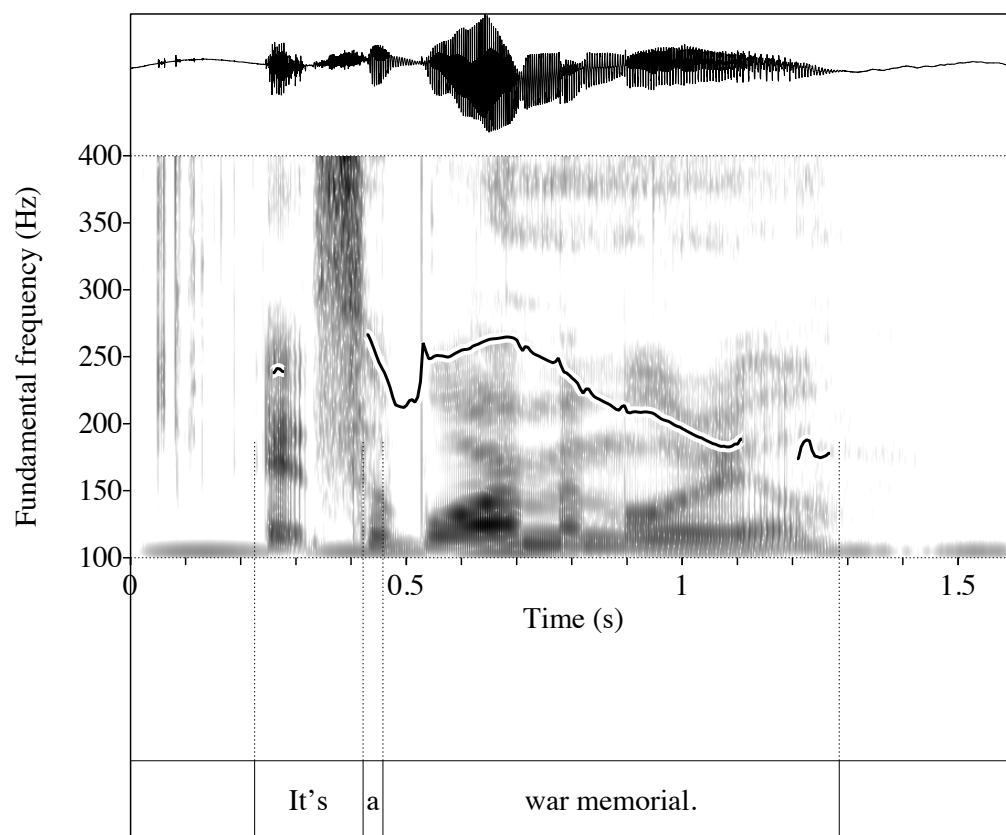
Figure 1.1: Spectrogram, waveform and $f0$ trace of the statement 'It's a war memorial' spoken by female Edinburgh English speaker SK. Notice the pitch falling across 'memorial'. This figure and all subsequent similar ones were produced using Praat (Boersma and Weenink, 2009)

Language change in lexical tone and lexical accent has been better studied than it has been in intonation, so it is possible that the study of intonational change

---

[3]For explanations of these terms, see appendix A.
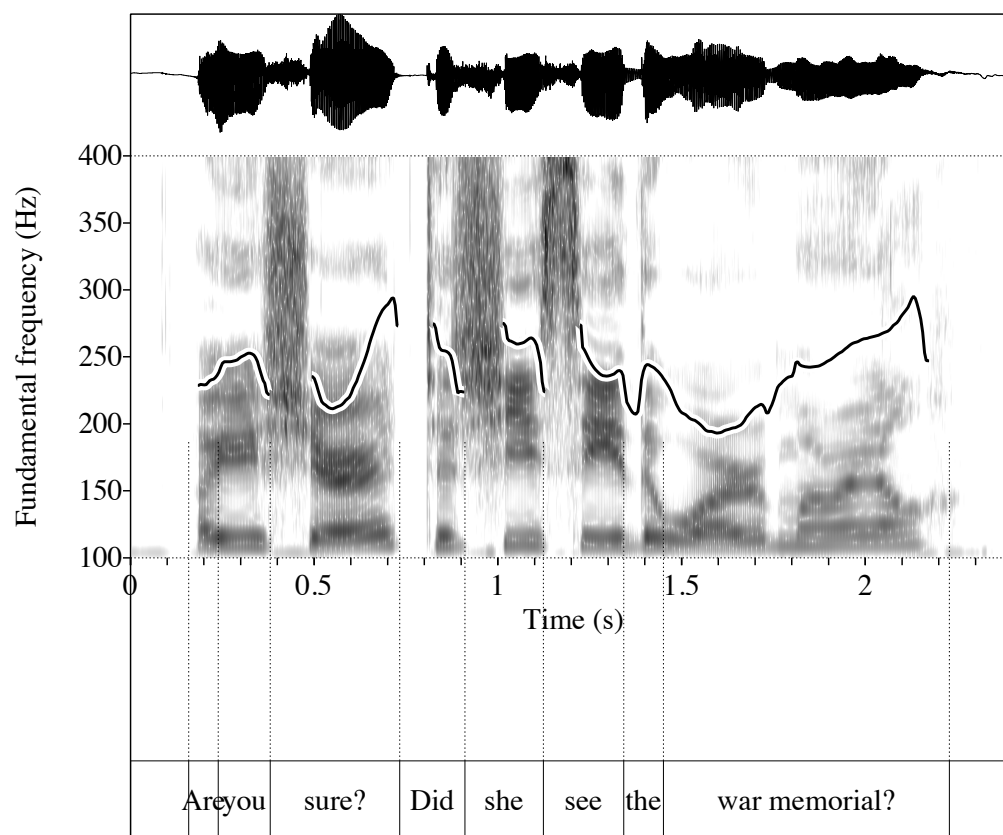
Figure 1.2: The questions 'Are you sure? Did she see the war memorial?' spoken by female speaker EL from Belfast. Notice the pitch dipping then rising across 'war memorial'.

may receive a boost from examining parallel phenomena that have been better studied in lexical tone and lexical accent.

1.3.2 Specific context in Belfast and Glasgow English

I begin to tackle intonational distance and change by focusing on a very narrow context: that of final statement intonational rises in Belfast and Glasgow English (Cruttenden, 1997; Ladd, 2008). Why do speakers in these varieties of English often produce a notable rise on stressed syllables on statements? Are these rises really 'rises' (Cruttenden, 1997)? Given that there are long-standing expectations that *statements* should contain falling pitch, and *questions*, rising pitch, this Belfast and Glasgow phenomenon has attracted attention for some time, but no clear-cut explanation has yet been found. Along with a number of other UK varieties,

Belfast and Glasgow have been included in a group of varieties termed 'Urban North British' (UNB) (Cruttenden, 1995, 1997; Ladd, 2008). There appear to be some remarkable similarities between UNB rises and statement rises in Donegal Irish,[4] which await a convincing explanation. This thesis does not involve any explicit comparison of data from the Irish and English languages. However, I make reference to Donegal Irish data as studied by Dalton (2007) and Dalton and Ní Chasaide (2005) on several occasions.

Figures 1.3 and 1.4 give an example each of a statement contour from Belfast and Glasgow English.
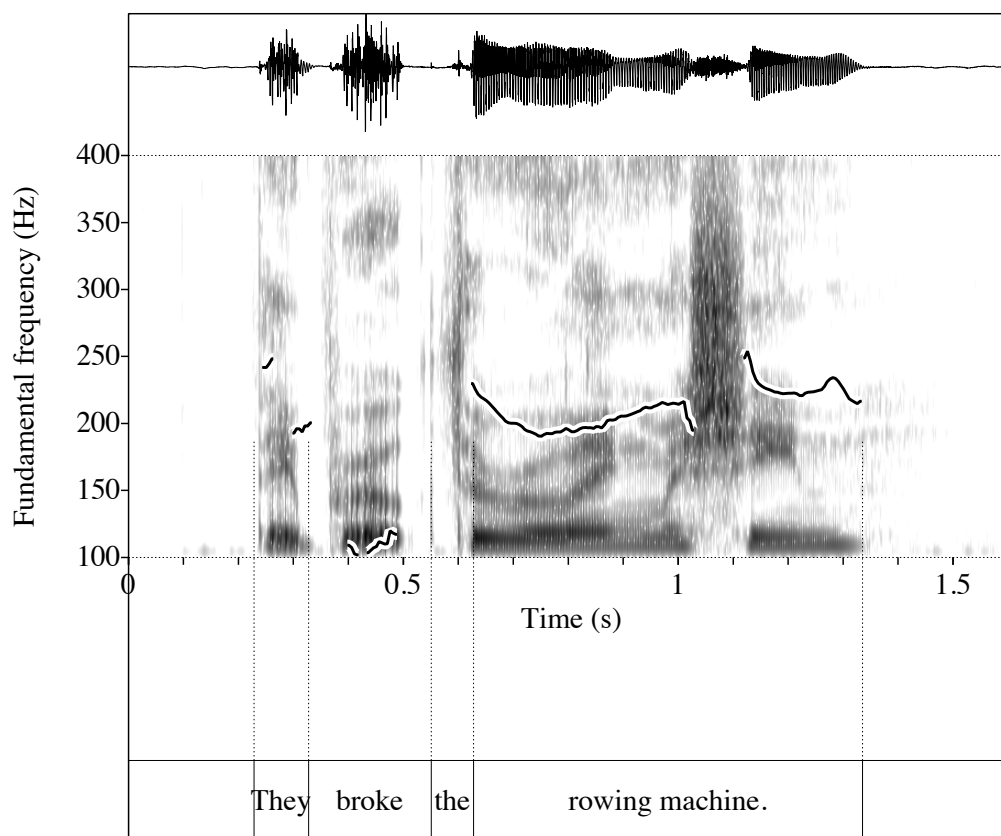
Figure 1.3: Spectrogram, waveform and *f*0 trace of a statement from female Belfast English speaker PT. Notice the rise beginning around the primary stressed syllable in the compound 'rowing machine'.

---

[4]Like Belfast, Donegal is located in the Ulster province in the north of the island of Ireland, though Donegal is part of the Republic of Ireland, whereas Belfast is in Northern Ireland.
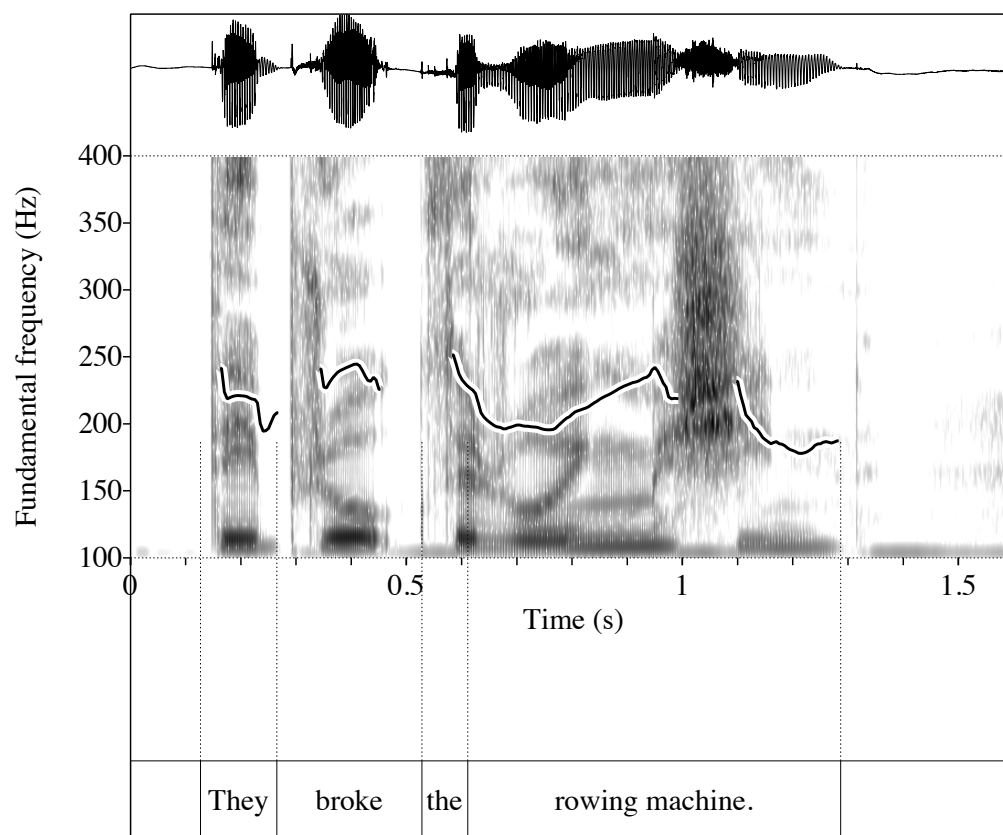
Figure 1.4: Spectrogram, waveform and *f*0 trace of the same statement, spoken by female Glasgow English speaker AG. There is also a rise on the stressed syllable in 'rowing'.

1.3.3 A fall changing into a rise

Could it be that originally, Belfast and Glasgow speakers produced the more usual final falls on statements and that over time, these final falls changed into rises? This is exactly the possibility I was referring to in my opening speculation that speakers of Belfast and Glasgow English may have once sounded more similar to speakers of other English varieties than they do today. Then, of course, the 'Where is the evidence' question follows on. The initial hypothesis is that if Belfast/Glasgow *statement* rises have derived from statement falls, then they would be more similar to statement *falls* than to *question rises*.

With phonetic measurements as intonational distances, it is possible to compare Belfast/Glasgow statement rises to statement falls in another English variety

and to question rises (either in Belfast/Glasgow or in another variety) and to do within-variety comparisons. This may allow the evaluation of whether the statement rises have come from statement falls or whether they might involve transfer of question intonation onto statements (Bolinger, 1978). These two possibilities become recast as the Alignment and Transfer hypotheses respectively in chapter 4.

These measurements crucially also allow the comparison of Belfast and Glasgow statement rises with each other. Cruttenden (1995) likens Belfast and Glasgow intonation to each other but admits that there is little experimental evidence to follow up these impressions. Their inclusion together in UNB assumes some sort of historical connection between them, perhaps due to the influence of speakers from the island of Ireland on English varieties or due to an "urban spread" (Cruttenden, 1995, 1997). Comparisons between Belfast and Glasgow allow me to ask whether Belfast and Glasgow statement intonation contours indeed have the same origin. They thus may help add to the understanding of potential intonational change. Hualde (2007) believes that studying synchronic variation among varieties of a language is a crucial part of working backwards in time.

1.3.4 Implications

The previous sections have provided an introduction to intonation and how different it is to segments. I have put forward the specific context in which I have chosen to begin approaching intonational distance and I have tried to demonstrate why this is a valid starting point. The two possibilities for the origin of the Belfast/Glasgow statement rises have implications for the relationship between change in intonation and change in segments. The Alignment hypothesis suggests that gradual phonetic change may account for the presence of rises on Belfast/Glasgow statements. This kind of change may have a counterpart in the theories of gradual phonetic change in segments (see for example McMahon, 1994; Blevins, 2004). The Transfer hypothesis adds a pragmatic dimension to sound change in intonation, which is not present in relation to phonetic change in individual segments. This is because the Transfer hypothesis posits the transfer of an intonation pattern with a specific utterance-level meaning (e.g. questioning meaning) onto a new domain, that of statements.

1.4 Fundamental questions

With the special context of Belfast/Glasgow statement rises now outlined, the fundamental research questions in this thesis are:

- What is the difference between attempting to measure distance in segments and in intonation? How could distance in intonation be measured?
- Can measurements of distance help in understanding how intonation might have changed?
- Is intonational change like segmental phonetic change?
- Do intonational distance measurements shed light on how to make progress in segmental distance measurements?

The central argument of this thesis is that intonation deserves a place within the context of the study of phonetic distance. Measurements of distance in intonation may be one tool to aid in understanding potential processes of historical change. Although studies of segmental change generally work at a much deeper time depth than is possible for intonation, phonetic distance measurements may be relevant to both of them. This thesis does not argue, however, that there is an absolute link between distance and historical change in intonation or segments.

Returning to the opening questions of this chapter (p.1), I hope it is now clear how these relate to the fundamental questions that I attempt to tackle.

- In exactly which respects did they sound similar?
- What exactly do I mean by 'similar' anyway?
- Where is the evidence?

'In exactly which respects' refers to 'melodic'/intonational distance as opposed to segmental distance. The step from segments to intonation represents the principal new direction in research for this thesis. 'What exactly do I mean by 'similar' ' refers to the approach of *quantifying* distance, away from intuitions. This chapter has already given an outline of the differences between existing distance measurements in segments and an appropriate way to begin measuring intonational distance. 'Where is the evidence' refers to the assessment of whether distance measurements can be used to explain potential processes of historical change (the origin of the Belfast/Glasgow statement rises) convincingly.

1.5 Contribution of the thesis

This thesis attempts to shed light on the following under-explored research topics:

- the treatment of *intonation* in the context of measurements of phonetic distance;
- mechanisms of sound change in intonation;
- an explicit comparison of the intonation of a few speakers of Belfast and Glasgow English;
- an assessment of whether intonation has any place in the phylogenetic frameworks increasingly used with distance measurements in segmental phonetics and in other areas of language;
- an exploration of how treatments of intonational distance may add to the understanding of segmental distance.

1.6 Overview of thesis content

I put forward how I see the progression of this thesis from the beginning to where it must end for the present.

1.6.1 Chapter 1

In this chapter, I have introduced the notion of phonetic distance, as a measurable concept. I have also made the crucial distinction between segments and intonation. It is the latter which represents the real direction of this thesis. Phonetic distance in segments has often been studied with historical questions in mind. This thesis will explore how far measurements of phonetic distance can help in understanding processes of historical change in *intonation* too. The specific context around which the thesis hangs is that of the phenomenon of statement rises in Belfast and Glasgow English.

1.6.2 Chapter 2

Since the concept of phonetic distance in intonation has been unduly left aside until the present, I ask how it has been tackled before in segments. I review and assess a number of existing segmental measures and expose some problems with them, making reference to a small-scale study comparing measures that I carried

out myself. The specific question I tackled in this small-scale study was how much phonetic detail should be in a measure, a question with relevance for intonation as well as segments. Phylogenetic techniques, originating in Evolutionary Biology, have found an ever-expanding second home in Historical linguistics. I touch on the positive benefits which these techniques have brought to measurements and also outline some problems with adapting them for language. I also raise the question of whether intonation may be appropriately analysed in any kind of phylogenetic framework. Although articulatory measures of phonetic distance dominate Historical linguistics and Dialectometry, I introduce acoustic segmental distances and I discuss their similarities and differences to articulatory measures and to my proposed intonational approach. I end the chapter by addressing why intonation has not been of interest to those devising phonetic distance measures.

1.6.3 Chapter 3

Chapter 3 looks specifically at intonation in the context of phonetic distance and sound change. I address why phonetic distances have been largely left to the side by those working on intonation and explore existing work on UNB varieties. I also introduce theoretical proposals on intonational change. These are in the context of the hypothesis that phonetic change has led to the development of statement rises in Belfast and Glasgow English from falls (recast as the Alignment hypothesis in chapter 4).

1.6.4 Chapter 4

Chapter 4 is an extensively experimental chapter. I give an overview of Experiment I, which used data from the IViE (Intonational Variation in English) corpus (Grabe et al., 2001), and involved comparing statement and question intonation in Belfast English and Cambridge English. I use two primary acoustic measurement parameters (Alignment, Scaling) as intonational distance measurements to evaluate the Alignment hypothesis. The results of Experiment I were problematic for this hypothesis. I also propose the Transfer hypothesis, that Belfast/Glasgow statement rises may involve transfer of rises typical either of questions or of continuation/list contexts. To evaluate the Alignment and Transfer hypotheses more thoroughly, I present the analysis of Experiment II. This involves a few Belfast and Glasgow English speakers and contains specially designed materials in two tasks:

- a Sentence reading task (including statements, lists, wh and yes/no (y/n) questions)
- Map task games (Anderson et al., 1991), designed to elicit more spontaneous speech.

The results from the alignment and scaling measurements show that Glasgow statement intonation supports the Alignment hypothesis while Belfast statement intonation supports the Transfer hypothesis. Thus, I argue that statement intonation inin these few Belfast and Glasgow English speakers respectively may have different origins.

1.6.5   Chapter 5

In chapter 5, I turn back to where I began, with measurements of phonetic distance in segments in the context of Historical linguistics. Does the intonational analysis make it possible to see this with fresh eyes? I explore how the individual intonational measurements of alignment and scaling may be incorporated into an overall intonational distance measure and into phylogenetic network techniques and what compromises must be accepted in this regard. The chapter concludes with what research on segmental phonetic distances could learn from intonation. I also expose the deep conflict between the needs of those using distance measures from phonetics (some of whom are non-linguists) and current trends in phonetics. I suggest how the two may be reconciled.

1.6.6   Chapter 6

Chapter 6 concludes the thesis. It shows why I have drawn a line in my research at this point, draws together the main themes and findings from the previous chapters and points out the best directions for future research.

1.7   Potential practical applications

Understanding the subtle phonetic issues which contribute to diversity and distance in varieties through systematic measurements is a valuable addition to knowledge in itself. A more thorough treatment of potential historical change in phonetics, particularly in intonation, where it has been so little studied, is also an important undertaking.

More generally, measurements of phonetic distance both in segments and in intonation may offer some practical benefit to those working in Forensic phonetics (Nolan, 1997; de Jong et al., 2007), Clinical linguistics (Connolly, 1997), Speech synthesis, Speech recognition, and 2nd Language learning (see Kessler (2005) in relation to segments).

For example, the act of quantifying distance along a set of phonetic features or parameters might enable the production of a scale on which a patient's improvement in pronunciation could be measured.

An isolation of some of the key phonetic differences in intonation between a few English varieties may help in the design of Text-to-Speech systems specifically for these varieties (Silverman and Pierrehumbert, 1990; Prieto et al., 1995; Rietveld and Gussenhoven, 1995; Kochanski and Shih, 2003; Dalton and Ní Chasaide, 2005; Arvaniti et al., 2006; Grabe et al., 2007).

Measurements between statement rises and question rises in Belfast/Glasgow English might be helpful in Speech recognition and 2nd Language learning to avoid the problem of statement rises being confused with questions (e.g. Haan, 2002; Grabe et al., 2005).[5]

1.8   Approaching Chapter 2

This thesis began with the distinction between words and 'melody'. More specifically, it takes as its starting point the measurements of phonetic distance that have been made for segments in individual *words*. I have briefly introduced the challenges of wishing to turn my attention from these segments in individual words to 'melody' (intonation), the latter needing very different types of measurement. Most pressingly, I need to show convincingly that there is no incongruency resulting from the leap that I am attempting to make from segments to intonation. Chapter 2 will start to address this by detailing how phonetic distance has been approached in segments (both with articulatory and acoustic methods), the increasing use of phylogenetic techniques in the field, and why intonation has not caught the attention of researchers in these areas until now.

---

[5]This problem involves pragmatic as well as phonetic issues.

CHAPTER 2

Comparing segmental distance measures

2.1 Introduction: Phonetic comparison among vowels and consonants

How similar is [o] to [ɑ]? Is [o] more similar to [ɑ] than to [i] or to a consonant such as [t]? These are questions about phonetic distance conceived at a basic level of individual vowels and consonants (segments). But why would a researcher wish to ask them? They might have historical interests, hypothesising that quantifying phonetic distance between the segments of particular words in different languages might indicate which of those languages are historically related. Alternatively, they might wish to study phonetic distance in the context of Dialectometry, establishing distances between pairs of varieties. Is there more segmental distance between Newcastle English and Dublin English than between Newcastle and Belfast English, for example? If the researcher could calculate a distance score, they might also be able to assess the degrees of intelligibility between speakers of closely related languages (Tang and van Heuven, 2007). There is also the setting of Clinical linguistics, where the researcher might like to assess how deviant a patient's pronunciation of a vowel or consonant is from the pronunciation by someone without a speech/language disorder (Connolly, 1997; Somers, 1999). These are just a few prominent domains in which distance measurements would be relevant (see Kessler (2005) for more). Speech synthesis and Speech recognition are two other fields in which segmental distance measurements have had great importance, but such measures have now been largely replaced with probabilistic techniques like Hidden Markov Models

(HMMs) (Holmes and Holmes, 1995; Coleman, 2005; Jurafsky and Martin, 2008). Therefore, I do not discuss these fields much further in this chapter.[1]

The next question to ask is 'From which perspective do I approach the development of segmental measurements'? There are three broad perspectives available:

- Articulatory (Heggarty, 2000a);
- Acoustic (Huckvale, 2004, 2007);
- Perceptual (Lindblom and Maddieson, 1988; Connolly, 1997).

2.1.1   Outline

This chapter focuses on the fields of Historical linguistics and Dialectometry in relation to segmental distance measures (sections 2.2 and 2.2.2). I show that there are strong connections between the approaches of researchers in these fields. The Edit Distance alongside *articulatory* phonetic feature systems have been the dominant tools (sections 2.3.2 and 2.4). Researchers have begun to incorporate their measures of phonetic distance into phylogenetic tree and network techniques (section 2.4.11). This represents a major trend within Historical linguistics and Dialectometry generally. The primary goal of this thesis is to move away from existing segmental distance measurements to phonetic distance measurements in intonation. The first step is to explore acoustic segmental distances (section 2.5) as intonation is also measured acoustically. Intonation has until now been of little interest to most historical linguists and dialectometrists. I conclude this chapter by examining why this has been the case (sections 2.6 and 2.7).

2.2   The role of Swadesh and Lexicostatistics

When I review the existing approaches to measuring phonetic distance among segments, one domain comes very much to the fore, Historical linguistics. In particular, the influence of Swadesh's technique of Lexicostatistics has been very strong (Swadesh, 1950, 1972). This technique in itself does not quantify phonetic distance but provides a measure of *lexical* distance between languages and varieties. In Lexicostatistics, this is achieved through a count of the number of shared cognates for a given set of meanings (e.g. the Swadesh 100 and 200 lists

---

[1]Nor do I make any reference to measuring distance in the context of the phonological framework of Optimality Theory (OT).

of basic meanings) between a pair of languages/varieties. The greater the number of cognates shared between a pair of languages/varieties, the closer they are deemed to be.

2.2.1 *Lexicostatistics becomes 'Phonostatistics'*

So what kind of relationship exists between Swadesh's Lexicostatistics and quantitative methods of phonetic comparison? The use of the term 'Phonostatistics' by Embleton (1986, 2000) rests on the assumption of a strong connection with Swadesh.[2] Heggarty (2000b) also makes specific links between his method of phonetic comparison and Lexicostatistics, specifically that phonetic quantification has the potential to be much less crude than simple cognate matching, extracting much more information from the data. Using Swadesh lists, Lexicostatistics involves taking a given meaning in two varieties and deciding whether the varieties are cognate or not for that meaning. So essentially it is a binary decision. For the meaning 'cold', German and English are cognate ('kalt', 'cold'), as both forms can be traced back by the linguistic Comparative method to a common ancestor (Joseph and Janda, 2003). However, for the meaning 'colour', German and English are not cognate ('Farbe', 'colour'), as the English word 'colour' is a borrowing from French (Stockwell and Minkova, 2001, p.38). 'Phonostatistics' also uses Swadesh-style meaning lists. The aim here is to work purely with cognates i.e. comparing the phonetic realisations of cognate forms in different varieties:

- RP [kʰəˑʊˑld]
- Standard German [kʰält] (McMahon et al., 2005-07).

So now the level of comparison is between individual segments. There are many different ways of making these segmental comparisons between varieties. These will be discussed in section 2.3 and section 2.4 below.

---

[2]However, when quantifying over lexical items, these items are usually independent of each other. Embleton (2000, p.160 note1) makes the important point that independence in phonetic data may be much harder to find than in lexical data. This has been hardly ever discussed or addressed, with the exception of Kessler (2001). It is clear that vowels and consonants are often not independent of surrounding vowels and consonants. It is beyond the scope of this thesis to deal with this problem comprehensively though I refer to it again at various points in relation to intonation in chapters 3-5.

2.2.2   Similar approach in Dialectometry

Though Swadesh's techniques were specifically developed to answer questions about historical relationships between varieties, his influence may still be felt in approaches to phonetic distance with the goal of purely synchronic comparison of varieties (Sullivan and McMahon, 2010). In Dialectometry, Séguy spearheaded the use of word/meaning lists of similar sizes to the Swadesh lists and made comparisons between transcriptions of these isolated words (Séguy, 1973; Goebl, 1984; Nerbonne et al., 1999; Heeringa, 2004). So there are direct parallels between Historical linguistics and Dialectometry in this regard.

2.2.3   Laver (1994)

Separately, Laver (1994) was one of the first to raise the importance of quantifying distance specifically within the domain of phonetics. He contrasted this need with the impressionistic judgements of distance which are implicit in standard phonological theories e.g. the well-known problem that English [h] and [ŋ], despite being in complementary distribution, are not phonetically similar enough to be considered allophones of the same phoneme.

2.2.4   Summary

In this introductory section, I have grounded existing approaches to measuring segmental phonetic distance in Historical linguistics and Dialectometry. I have shown that researchers in both fields have taken a broadly similar approach to each other, using meaning lists such as the Swadesh lists, which were first used in lexical comparison. They have used individual transcriptions of these list items to make comparisons between vowels and consonants between different languages/varieties. In section 2.3 and section 2.4 below, I examine some of the existing measures in greater detail.

2.3   Assessment of measures and problems exposed

2.3.1   Outlining the context

My central question in the assessment of different measures of phonetic distance developed for segments is how 'phonetic' they actually are. I define phonetics

as the study of speech sounds, their articulation through the changing configurations of the vocal tract, their acoustic representations, and their perception (Ladefoged, 1975; Ladefoged and Maddieson, 1996; Hardcastle and Laver, 1997; Hardcastle et al., 2010).

I first draw attention to the constraints already in place. The use of meaning list items like those in the Swadesh 100 and 200 lists presupposes phonetic comparison between transcribed vowels and consonants in words occurring in isolated as opposed to connected speech. The major advantage of this is that it easily addresses the latter of the two problems identified by Heggarty (2000a) as the Quantification and Compatibility problems (see section 2.3.4 below). The Compatibility problem grapples with how to decide which words and which segments should be compared with which. Comparing cognate forms (e.g. English 'cold' against German 'kalt') is based on a robust conception of comparing like with like.

Isolated word comparison is acceptable in historical linguistic contexts because often other kinds of data are simply not available. Dialectometrists would be in a position to quantify more than mere word pronunciations. Therefore, using only isolated word transcriptions with contemporary data is somewhat stilted. However, the background of the influence of Swadesh's techniques on Dialectometry as well as on Historical linguistics gives a plausible explanation as to why many have remained within this straitjacket. Laver (1994) is one of the few to outline how phonetic distance could be measured in connected speech.

2.3.2   Edit Distance

One of the most basic comparisons that can be made between a pair of cognates is the Edit/Levenshtein distance between them (Levenshtein, 1966; Wagner and Fischer, 1974; Heeringa, 2004).

The Edit Distance algorithm compares computational strings. In a phonetic context, the strings are comprised of phonetic symbols. This algorithm finds the least costly way of transforming one phonetic transcription into another e.g. RP and U.S. English pronunciations of 'home':[3]

- RP [həˑʊˑm]

---

[3]These are IPA transcriptions of individual speakers deemed to be representative of the respective varieties of English.

- Standard U.S. English [hʌ̟·ʊ·m] (McMahon et al., 2005-07).

Transforming one transcription into another involves insertions, deletions or substitutions of phonetic symbols. The least costly transformation is the one that uses the least insertions (I), deletions (D), or substitutions (S).[4]

Edit Distance

$$D(r,t) = min(I + D + S)^5 \tag{2.2}$$

The use of the Edit Distance here is akin to Lexicostatistics in its binary nature. Lexicostatistics assesses whether two words are cognate or not. The Edit Distance assesses whether two phonetic symbols are identical or not. If they are identical there is no cost. If they are not, they receive the universal cost of 1 e.g. comparing [o] against [a] would receive the same cost as [o] against [g]. Converting the U.S. English transcription of 'home' above into the RP transcription would just require one substitution of the vowel sounds and no insertions or deletions. It is clear that treating speech sounds in this way entails treating them as symbolic, discrete, letter-like elements.

The Edit Distance also links with Lexicostatistics in another way. Deciding whether two forms/words in two different languages are cognate is not as straightforward as it might appear from the pair 'cold' and 'kalt' (section 2.2.1 above). It is a subjective process, using the long-established linguistic Comparative method in Historical linguistics (Joseph and Janda, 2003; McMahon and McMahon, 2005). Since cognate pairs are usually more phonetically similar to each other than to other non-related forms, researchers have used the Edit Distance (and versions of it) as an objective alternative to subjective cognate judgements (Bakker et al., 2009).[6]

---

[4]Henceforth, I use the term Edit Distance as opposed to Levenshtein distance. The two terms are often interchangeable. Strictly speaking, the Levenshtein distance assigns a single point for each insertion, deletion, or substitution. The Edit Distance may apply different costs/weights (W) to the insertions, deletions and substitutions.

[5]Weighted Edit Distance

$$D(r,t) = min(W_1 I + W_2 D + W_3 S) \tag{2.2}$$

[6]Campbell (2003) shows that there can be cognate forms which are not phonetically similar though.

The most extensive work using the Edit Distance has been in the field of Dialectometry rather than in Historical linguistics (Kessler, 1995; Heeringa, 2004; Nerbonne and Heeringa, 2010). These researchers and their colleagues have argued quite strongly for its adequacy in dialect comparison, though Heggarty (2006) has severely criticised its crudity. The computational simplicity of the Edit Distance is a strong attraction, researchers acknowledge (Kessler, 1995; Heeringa, 2004; Bakker et al., 2009). The tension between computational simplicity and phonetic faithfulness is evident throughout attempts to measure phonetic distance.

Although it is prominent in the context of phonetic distance measurements, the Edit Distance does not strictly fall into any of the categories of articulatory, acoustic, or perceptual distance outlined in section 2.1 above. In fact, it disregards articulatory (or other) proximity of sounds entirely; it is only interested in the binary distinction between identical and non-identical letter-like symbols.

In addition, do Edit Distances really reflect segmental comparison at all? The use of isolated word transcriptions results in an Edit Distance score for a pair of these transcriptions. The pair of transcriptions of 'home' above would simply score 1. It is then standard to normalise for word length, but this normalised score is immediately at one remove from the segmental comparisons themselves. It could be argued that the use of isolated words combined with the use of the Edit Distance algorithm actually results in word-level comparison and not segment level comparison (Heggarty et al., 2005; McMahon et al., 2007). This would further dilute the phonetic component of this kind of comparison. For example, comparisons between the same two segments could end up with different scores depending on the size of the words in which they were embedded.[7]

2.3.3 'Matching' versus articulatory gradation

The most fundamental problem with the binary nature of the Edit Distance and with a concept of 'match' vs 'non-match' is that it relies on discreteness and does not allow any kind of articulatory gradation.

Moving away from these binary values of Lexicostatistics and Edit Distances, it is possible to attach numerical values to various phonetic feature systems (e.g. Delattre, 1965; Jakobson et al., 1952; Ladefoged, 1975; Ladefoged and Maddieson,

---

[7]Heggarty (2000a) admirably attempts to steer clear of word level comparison in his own method.

1996; IPA, 1999). These break sounds down into articulatory or other dimensions and thus can express degrees of distance along these dimensions. However, within these methods, many levels of subtlety exist. These feature values may still be used in comparisons of cognates in Swadesh-style lists. For further details, see section 2.4 below. So I envisage a kind of cline extending from the kind of binary cognacy judgements of Swadesh's Lexicostatistics, to the binary Edit Distance application within cognates, to the binary assessments of initial consonant matches, to the more graded feature methods. The Edit Distance along with articulatory feature methods have been dominant over acoustic and perceptual methods in approaches to measuring phonetic distance thus far, though acoustic measures are increasing (section 2.5 below).

2.3.4   Heggarty (2000a)

At the most phonetically explicit end of the cline from Lexicostatistics, through basic Edit Distances, to phonetic feature systems is the method of Heggarty (2000a). Heggarty's method is predominantly articulatory. It rests on the tenets of the IPA and Laver (1994). One of Heggarty's most important contributions is his clear exposure of two fundamental problems facing any attempt to measure similarity/distance in phonetics.[8] These are the Compatibility and Quantification problems. Heggarty is the only researcher who has addressed the former thoroughly. Before measuring, it is imperative to have a concept of comparing like with like. Heggarty deals with this strictly through the use of cognates and attested ancestral forms. The ancestral forms are used as templates and broken into slots as are the cognates themselves. Existing knowledge of the sound changes that have occurred enable appropriate slots in the cognates to be compared with each other. Heggarty assigns a particular phonetic length to each slot (Heggarty et al., 2005, p.58). It is possible for one slot in a cognate to be compared against two or more slots in another cognate (McMahon et al., 2007, p.120). The basic unit is the phone (Laver, 1994, p.571-2) with a single articulation and standard length (McMahon et al., 2007, p.125). This is a key difference with the methods of Heeringa (2004) in which only single individual segments may be compared with each other (section 2.4 below).

The Quantification problem deals with how to measure distance between corresponding slots. The details are extremely complicated in Heggarty's method

---

[8]Heggarty himself refers to his method as a method of phonetic *similarity*, which is why I have re-introduced the term 'similarity' here.

and are not thoroughly explained or illustrated in any of Heggarty's publications. However, the principles upon which the numerical values are derived are made explicit:

- Natural Principles of Language Structure,
- Cross-linguistic Norms, and
- Logical Principles (Heggarty, 2000b, p.541).

Heggarty's method takes into account not just segments but a sub-level of gestures (Browman and Goldstein, 1992), secondary articulations and the related timing phenomena that accompany them (Heggarty et al., 2005, p.51). However, he does so within the strictures of the IPA system and its diacritics. These fine-tuned differences are deemed to be particularly necessary for the study of closely-related varieties (McMahon and McMahon, 2005, p.225) and indeed Heggarty and his colleagues have done more than other historically oriented scholars to incorporate both historical and contemporary varieties into quantitative phonetic comparison (McMahon et al., 2005-07, 2007; Maguire et al., 2010).

Heggarty justifies his method entirely on these theory-internal principles (Heggarty, 2000a, p.42), unlike others who look to external measures of 'success' such as perceptual judgements (Heeringa, 2004) or existing language classifications (Kessler, 2007).

2.3.5 The problem of phonemic contrasts

One of Heggarty's principles that is especially worthy of further discussion is the Cross-linguistic norms principle. The fact that most languages display a binary phonemic contrast in voicing as opposed to a three-way contrast in place and manner of articulation is taken into account in the weighting process (see Heggarty (2000a, p.152), Heggarty et al. (2005, p.52) and McMahon et al. (2007, p.119)).[9] It is important to understand exactly the relevance of such *phonemic* distinctions in a method of *phonetic* similarity/distance especially since it is of great importance to Heggarty to keep the distinction between phonetics and phonology clear. Heggarty (2000b, p.543) justifies his use of these common phonemic distinctions by assessing them as "a cross-linguistically balanced evaluation of the 'neutral' relative significance of incremental phonetic differences". However, the use of phonemic distinctions may not be the most appropriate way to set

---

[9]By 'weighting', I mean the different numerical values that are assigned to feature contrasts.

up such a baseline. One reason is that there are segmental phonetic differences between languages and varieties that are not related to typical phonological contrasts (Ladefoged and Maddieson, 1996, p.5) and thus even in Heggarty's detailed method, some of these may be missed.

It is also useful at this point to refer to Ellison and Kirby (2006) and Port and Leary (2005), who claim that phonetic 'space' is moulded differently for different languages. On this basis, Ellison & Kirby argue that phonetic distance may only be measured within a single language and not cross-linguistically. Their view could not happily co-exist with Heggarty's incorporation of Cross-linguistic norms into his model. The example given in support of Ellison & Kirby's claim is that there are cases in languages in which sounds considered very different from an articulatory point of view are in fact treated in that language as homophones (see Ellison and Kirby (2006, p.274) on the apical and uvular trill in European Portuguese). This point shows how fine the distinction between phonetic and phonemic actually is.

Heggarty (2000a, p.72) also admits that his time-slots "should be seen not so much as universal but as realisations of language-specific phonological distinctions, i.e. phonemes". So the problem of the reliance on discrete elements (somewhat) removed from articulatory gradations re-appears, even in the midst of Heggarty's attempts to incorporate the sub-phone level.

Heggarty has argued that his method is superior to the Edit Distance, simply because it takes far more phonetic detail into account (e.g. Heggarty, 2006). Nevertheless, his faith in the appropriateness of feature systems for phonetic comparison is somewhat at odds with the most recent trends in phonetics where researchers are increasingly critical of the validity of features, phones and the treatment of speech as a linear sequence of discrete segments (Port and Leary, 2005; Ladd, forthcoming). Heggarty's method does indeed incorporate the gestural level but still within the context of the IPA diacritics, so he is still subject to the IPA assumptions (IPA, 1999, p.5) about the discretisation of speech.

In chapter 5, I re-assess some of the points I have made about Heggarty's method.

2.3.6 *Initial consonant comparisons*

Another method has simply compared the initial consonant of corresponding forms in different languages. But why omit phonetic comparisons from later

portions of corresponding forms? Kessler has argued strongly for this minimal approach in his own work (Kessler, 2001; Kessler and Lehtonen, 2006; Kessler, 2007). He uses these measurements to test the hypothesis that two or more languages are related. He argues for comparing only those phonetic elements least likely to change over time and therefore most likely to represent an ancestral connection between two languages. In the Indo-European languages, the place of articulation of the 1st consonant is known to be particularly stable. Kessler (2001) is convincingly able to show that comparisons of other parts of the words give scores which obscure relationships between languages that are known to be related. However, Heggarty et al. (2010) argue that Heggarty's own method (which takes far more into account than characteristics of the initial consonant) works well at showing historical connections in practice. The initial consonant has also been prominent in attempts to define a level of chance phonetic distance between a pair of languages. The goal here is also to establish or confirm that a pair of languages with a significantly smaller distance score than this chance level are related (Oswalt, 1970; Ringe, 1992; Baxter and Manaster-Ramer, 2000; Kessler, 2001).

2.3.7 *Previous comparisons of measures*

There is a scarcity of studies comparing different measures of phonetic distance, though Heeringa (2004), Kondrak and Sherif (2006) and Kessler (2007) are exceptions. What these studies have in common is their use of phonetic distance measures for non-phonetic purposes: correlation with overall listener judgements of distance between varieties (Heeringa, 2004), or validation through the measure's ability to identify known cognates (Kondrak and Sherif, 2006) or known language relationships (Kessler, 2007). By contrast, no one had examined whether the phonetic differences encoded in different measures actually affect the overall distance scores between pairs of varieties. This became the goal of my own study (section 2.4 below).

Phonetic feature methods intrinsically carry much more phonetic detail than the basic Edit Distance, yet Heeringa (2004) found very similar correlations with listener judgements for feature methods and basic Edit Distance. This convergence has been unexpected to the researchers themselves (Nerbonne and Heeringa, 1997). Different feature systems do not treat each articulatory contrast in identical ways. Yet again, Heeringa (2004) found very similar correlations between listener judgements and all of the feature methods he examined. Kessler (2007)

and Kondrak and Sherif (2006) also found the different methods they used performing almost equally as well as each other on identifying known language relationships and identifying cognates.[10] This suggests that somehow the range of differences in these methods get levelled out. So it is clearly not a straightforward case of the more features the better!

> Given the huge number of parameter combinations that may be invoked in defining distance measures, the danger looms that researchers are merely industrious, not insightful in developing techniques (Nerbonne, 2007, p.432).

However, the convergence may also be a consequence of these researchers methods of evaluation. The goals of comparison against listener judgements, cognate identification and constructing known language relationships may be far too broad to test the subtle phonetic differences between feature systems appropriately. In fact, phonetics may not actually be necessary for any of these goals. Indeed, the reason why the Edit Distance may indeed appear to be adequate for the goals it has been set so far is that these goals do not really test its phonetic remit. It is thus unsurprising that comparison essentially at the orthographic level would appear sufficient.

I have indicated that the few previous studies that have compared different measures tend to show that very different methods result in very similar patterns of connections emerging between varieties. One could use this as confirmation of the particular variety groupings found. Indeed, some other studies (not exclusively phonetic studies) do adopt such an approach (Embleton, 1986; Nakhleh, Warnow, Ringe and Evans, 2005; Moisl and Maguire, 2008).

Another way of approaching this convergence is to question what those groupings are actually based on if they will emerge from almost any method. If different phonetic feature methods produce a very similar result to each other and to a cruder comparison like the basic Edit Distance, then it is clear that subtle phonetic differences are not at the root of the connections between these languages/varieties. Rather it must be a much broader level of connection e.g. broadly phonemic. If "numerical approaches can reveal patterns which are real

---

[10]However, Kondrak & Sherif did find Kondrak's own feature method ALIGN (Kondrak, 2000) performing better at cognate identification than letter-like Edit Distance based measures.

but marginal" (McMahon et al., 2007, p.116), it is important to access these patterns. To understand how phonetic (and other) distances really work, it is necessary to look closer at the target areas in which the feature methods are different from each other (Warnow, 2009).

2.4 Small-scale study

2.4.1 Introduction

I compared the Edit Distance and 3 articulatory feature methods on the same data, a small subset of 30 meanings from the 110 meaning list of purely Germanic cognates compiled by McMahon et al. (2005-07). I chose a subset of five Germanic varieties initially: Standard American English, RP English, Standard Scottish English, Buckie (a variety of North-Eastern Scots) and Standard High German. Later, I added five more varieties: Liverpool English, Tyneside English (typical), Belfast English, Norwegian (Stavanger), and Swedish (Stockholm).

2.4.2 Phonetic feature systems used

The feature systems I used with numerical values already attached were as follows:

- the version of the Almeida and Braun (1986) method used by Heeringa (2004, p.40ff) and Heeringa and Braun (2003) (hereafter 'Heeringa version');
- the original Almeida and Braun (1986) method (designed for measuring intertranscriber reliability), which is quite different from Heeringa's use of it;
- the Connolly (1997, p.282ff) method (which is based on Ladefoged (1975)).[11]

I adapted each method slightly for use with my particular data set.

One important omission to my initial comparison of feature methods was the method of Heggarty (2000a). There were two primary reasons I excluded Heggarty's method from empirical testing:

- Implementing his method alongside the other measurements would have required the data to be in a very different format and thus comparisons between methods would not be straightforward;

---

[11]Also used and adapted by Somers (1999) and Kondrak (2000).

- Inclusion of his method was not essential to my overall goal of moving from segmental phonetic distance to intonational distance.

I also omitted acoustic segmental measures from this intitial comparison. To include such measures would also have required substantial extra data analyses (e.g. measuring formant values instead of simply assigning articulatory features). Even though acoustic segmental measures may appear on the surface to have more in common with the acoustic intonational measurements that I proceed to in later chapters, there was not enough overlap between acoustic segmental measures and my proposed intonational approach to merit the former's inclusion in this exploratory comparison of segmental measures. I do discuss acoustic segmental measures in section 2.5 below, however.

2.4.3 *Some differences between the measures*

The 'Heeringa version' is based on the number of IPA categories for the standard 3-way classification of consonants and separate 3-way classification of vowels. One of the negative consequences of this approach is that a difference between a high and a low vowel incurs a much greater distance than the difference between a back and front vowel or a rounded and unrounded vowel, simply because the IPA makes more distinctions on the height axis than elsewhere. This is not as much of a problem in the original Almeida & Braun approach. Rather than use the full range of IPA distinctions, Almeida & Braun constrained their scoring focussing on neighbouring distinctions. Greater differences along a parameter are not distinguished by numerical values, but are rather given the same ceiling score. For example, among vowels, differences of two or more levels of height are uniformly given a score of 2.

The main differences between the Connolly method and the 'Heeringa version' are Connolly's use of several different features as opposed to just one to represent consonantal manner of articulation, vowel roundness incurring a higher cost in Connolly, and fewer vowel height distinctions in Connolly. Full details of the phonetic features and their numerical values (weightings) for each method are in appendix B.

Heggarty's method has the advantage over the other methods, in principle, of having a motivation for each numerical score assigned to feature contrasts. The distances the other methods assign to particular vowel and consonant contrasts

and the relative weighting of these distances against each other are very much implicit and unjustified.[12]

2.4.4 Distances between individual segments

The first step is to draw up a vector of feature values for each segment. I did this using Microsoft Excel and devised my own numerical codes for each vowel and consonant. For example, the 'Heeringa version' uses scales from 1-3, 1-7 and 0-1 for vowel backness, height and rounding respectively.[13] The vowel [a] would have the following bundle of numerical values:

$$\begin{bmatrix} 1(front) \\ 7(open) \\ 0(unrounded) \end{bmatrix}$$

The vowel [u] would have these values:

$$\begin{bmatrix} 3(back) \\ 1(close) \\ 1(rounded) \end{bmatrix}$$

The next step is to compare vectors for pairs of segments. To do this, I used the Manhattan ('City block') distance for two of the feature methods ('Heeringa version', Connolly). Values for corresponding features are subtracted from each other. The resulting values are then summed to give a score of distance between two segments.[14]

$$d(p, q) = \sum_{i=1}^{n} |p - q|_i \qquad (2.3)$$

---

[12]Kondrak (2000) also provides some justification for the feature scores in his measure. He sharply points out the problem that the comparison of [p] and [pʰ] in the Connolly system scores a greater distance than the distance between [p] and [k].

[13]Although the IPA transcriptions of the data from McMahon et al. (2005-07) are narrow with the use of the full range of diacritics, I only incorporated diacritics into the scoring procedure at the earlier stages of the analysis. I compared the results of incorporating the diacritics and omitting them and there was no difference. This indicates that small differences may make no impact.

[14]Connolly (1997) is an excellent, clearly-written article explaining details of feature matching, the Manhattan distance and other distances.

For example, the Manhattan distance between [a] and [u] in the 'Heeringa version' would be as follows:

$$|(1-3)|(backness) + |(7-1)|(height) + |(0-1)|(roundness) = 9 \qquad (2.4)$$

I used the R program (R Development Core Team, 2010) to calculate Manhattan distances between all pairs of consonants and vowels respectively in both of the above feature methods (using the **dist** function). The Manhattan distance is a commonly used metric for computing distances between segments (Grimes and Agard, 1959; Connolly, 1997; Heeringa, 2004; Kessler, 2005).

I did not use the Manhattan distance with the original Almeida & Braun method. This is because Almeida and Braun (1986) do not apply a straightforward bundle of features to each segment. Rather they apply ceiling scores of 2 and 3 between many pairs of segments. For details, see appendix B.

Using Manhattan distances (Heeringa, Connolly) and Almeida & Braun's own values, I produced a matrix of pairwise distances between all vowels and consonants respectively in each system.

2.4.5   Distances at the 'word' level

The next step is to use these segmental distances to compute a distance score between a pair of transcriptions of a cognate word/meaning. To calculate the word distances using the Edit Distance and using the feature systems, I used the **sdists** function in the cba package in R. The **sdists** function produces a matrix of distances between varieties for each word separately. The input was a separate comma-separated spreadsheet file for each of the 30 words. This contained the transcription of that word in each variety arranged in columns.

2.4.6   Basic Edit Distance

For the basic Edit Distance, I used the **sdists** function with weights set to 1 for insertions, deletions and substitutions and 0 for matches of segments.

I then normalised for differences in word length, by dividing each pair-wise score by the length of the longer transcription in number of phonetic symbols. This is a standard approach (Serva and Petroni, 2008; Bakker et al., 2009).

As an example, I use the transcriptions of the numeral 'eight' in Buckie and Std Scottish English (SSE).

- Buckie [ä·χt]
- SSE [et^h]

The basic Edit Distance is 2. Converting the Buckie transcription into the SSE one requires a substitution of the initial vowel and the deletion of the consonant [χ]. The Buckie transcription is the longer one, consisting of three segments. The normalised Edit Distance is 2/3 = 0.67.

2.4.7 Feature methods

For the feature methods, I adopted elements of the framework of Heeringa (2004). Heeringa presents the feature methods as also being implemented within the Edit Distance algorithm. However, the computation of word distances in the feature methods and basic Edit Distance respectively actually have some strong differences.

When comparing two transcriptions of a particular word/meaning, corresponding vowels and consonants must be lined up. This addresses Heggarty's Compatibility problem, the problem of comparing like with like (section 2.3.4). To get a distance score at the 'word' level, the distances between pairs of corresponding vowels and consonants are summed. This is different from the basic Edit Distance in the following main respects:

- More gradient distances reflecting articulatory proximity are possible in the feature methods. For example, in the basic Edit Distance, [i] vs. [e] scores the same as [i] vs. [a], but the latter pair incurs a greater distance in the feature methods.
- The basic Edit Distance algorithm does not decide in advance which elements should be compared with which. Rather, the Edit Distance finds the least costly alignment of segments. In the feature methods though, I only allowed vowels to be compared with vowels and consonants with consonants in this initial study, largely following Heeringa.

For the basic Edit Distance, it was not a problem that the transcriptions of a given word did not have equal numbers of symbols. However, for calculating distances with feature values while using the **sdists** function, it was necessary

that all transcriptions be of the same length. To do this I lined up corresponding consonants and vowels by hand, inserting a special code to represent 'silence' if a segment in one variety was without a counterpart in another variety (see section 2.4.8 below).

For the feature methods, I used the **sdists** function in the 'alphabetic weights version'. The weights were comma-separated files of the matrices of pair-wise segmental distances for each feature method.

2.4.8 *Presence vs. absence of segments*

Returning to the issue of 'silence', what if a consonant/vowel present in the pronunciation of a word in one variety was absent in another variety e.g. in comparing rhotic and non-rhotic varieties of English? Heeringa (2004, p.71) deals with this problem by choosing [ʔ] and [ə] to represent 'silence' in consonants and vowels respectively. I initially followed this approach but there are several problems with it. For example, /r/ is produced with very different articulatory mechanisms to [ʔ], so Heeringa's approach to 'silence' results in large distances automatically emerging for the rhotic/non-rhotic divide in English varieties. Heeringa (2004, p.79-80) acknowledges that in feature systems, defining 'silence' is not easy and is 'artificial' but claims that it is necessary for the Edit Distance framework in which he bases his work. The choice of [ʔ] and [ə] for 'silence' in consonants and vowels may be compatible with maximally unmarked segments phonologically (McMahon, 2000). In practice though, there is not an obvious phonological/phonetic justification for why inserting [ɹ] should incur a higher distance than inserting [ʃ], simply because the latter is closer to [ʔ] than the former. Heggarty avoids 'silence' by allotting a specific phonetic length to each segmental slot in the ancestral form (Heggarty, 2000a; Heggarty et al., 2005). The positive aspect of this is that one segment can be compared against part of another segment, reflecting language changes like compensatory vowel lengthening following consonant loss. Unfortunately, it was not possible to implement this with the **sdists** function in R.

Following Connolly (1997, p.291), for his system, I chose a fixed cost of 1 for the presence/absence of a segment (Kondrak, 2000; Kondrak and Sherif, 2006) and re-scaled all the other costs between pairs of segments so that they were no more than 0.8. With the Almeida & Braun original system, I simply applied the ceiling of their scales for vowels and consonants to the presence/absence of

segments too. This is because they only make a broad distinction between the neighbouring differences between sounds and all other bigger differences.

Admittedly, any treatment of 'silence' makes phonetic comparison like the presence or absence of typed keyboard letters. In historical sound change, consonants or vowels rarely if ever are lost or gained without important effects on the surrounding vowels or consonants in the words affected. Kessler (2005) observes that the Edit Distance algorithm can take no account of the surrounding phonetic environment to a given segment.

2.4.9   Illustration of feature methods

As an illustration, I provide examples of how to calculate word distances between the Buckie and SSE transcriptions of 'eight' using the 'Heeringa version' and the Connolly method.

The distance in the 'Heeringa version' is calculated as follows:

- [ɑ] (open back) vs. [e] (close-mid front) scores 6 (close-mid vs. open 4, back vs. front 2).
- The vowel length difference scores 0.5.
- I did not incorporate aspiration in the Heeringa version so [t] vs. [tʰ] scores 0.
- [χ] vs. 'silence' (essentially [ʔ]) scores 7 (stop vs. fricative 4, uvular vs glottal 2, consonant vs. not consonant 1).

Adding all these scores together gives a total score of 13.5. Dividing by 3, the length of the longer transcription, gives a normalised score of 4.5.

The distance in the Connolly method is calculated as follows:

- [ɑ] vs. [e] scores 1.5 (1 for the backness difference, 0.5 for the height difference).
- The vowel length difference scores 0.1. The total vowel score is thus 1.6.
- [t] vs. [tʰ] scores 0.5.
- [χ] vs. 'silence' scores 1.
- Re-scaling the other scores by 0.8 gives 1.28 for the vowel difference and 0.4 for the aspiration difference.

This gives a total score of 2.68. Dividing by 3 gives a normalised score of 0.893.

See appendix B for further details on any of the individual segmental scores.

2.4.10 *Overall scores*

The previous sections have explained how phonetic distances in the Basic Edit Distance and in the feature methods are calculated at the level of the individual words. Once I had scores for each word between each pair of varieties, I calculated the mean distance between each pair of varieties over all words for each method. Means and standard deviations are in appendix B.

It is important to identify the two levels here, clearly outlined by Nerbonne et al. (1999, p.viii):

- the level of individual segmental comparisons in word transcriptions;
- the level of variety comparisons, which are aggregates of these individual segmental comparisons.

The use of an aggregate score has generally been accepted without question. Rarely do researchers query whether their data is normally distributed enough for a simple aggregate to be appropriate.

Separately, the process of computing distances and aggregating them is well-acknowledged to result in a great loss of information (Embleton, 2000; Holder and Lewis, 2003). When sounds are broken down into features and when distances are computed over a range of sounds in a word, it may be difficult to ascertain exactly if and how the detail in the features is affecting the overall score.

2.4.11 *Phylogenetic Network techniques*

Historical linguists have been influenced by the explosion of new techniques in Evolutionary Biology. These techniques, broadly termed Phylogenetic techniques, are designed to study processes of vertical and horizontal transmission in biological taxa. Within this group of techniques, there are two further basic distinctions:

- between Distance- and Character-based methods,
- between Tree and Network methods.

For a clear overview of these techniques, see McMahon and McMahon (2005) and Renfrew and Forster (2006).

In a language context, distance-based methods take a matrix of distance scores between a set of languages/varieties as their input (e.g. the aggregate phonetic distance scores outlined in section 2.4.10 above). These distances are used to pair up languages/varieties and build a tree or network using clustering algorithms. Character-based methods use a small set of Characters, akin to sociolinguistic variables (McMahon and McMahon, 2005), with discrete values called Character states. Character values are not added together to give an overall score, unlike in distance methods. Bandelt et al. (1999) discusses one algorithm for combining the Character values. Heggarty (2006) argues that linguistic data are not as easily measured as biological data nor as discrete so it is still an open question how appropriate these methods are for linguistic data.

Since phonetic distances computed over words consist of incremental numerical scores, it is straightforward to use them with distance-based methods. Would a Character approach be possible? I attempted this with the character method 'Network' (Bandelt et al., 1999). The complications included the restriction that only binary character formats could be used with non-genetic data. This prevents articulatory gradations being easily expressed and makes for cruder assessments of identity and non-identity of segments. However, Heggarty et al. (2010) has made more inroads using 'Network'[15] with Heggarty's method of phonetic similarity/distance.

Network methods are intended to display processes other than tree-like descent. If the data do not unambiguously suggest a tree structure, network methods produce multiple trees on top of each other. The branches of these multiple trees are connected by reticulations.

With phonetic data, NeighborNet (Splitstree) (Huson and Bryant, 2006), a distance-based method, has been extensively used by McMahon, Heggarty, and colleagues. NeighborNet uses a modified version of the Neighbor Joining Algorithm (Bryant and Moulton, 2002; McMahon et al., 2007). I also used Neighbor-Net with the segmental phonetic data I examined and present the results below (section 2.4.12). Phylogenetic networks are not the only tools that have been used with phonetic distance measurements. Prokic and Nerbonne (2009) are the

---

[15]'Network', when appearing inside single quotations like this refers specifically to the programme by Bandelt et al and not to network methods generally.
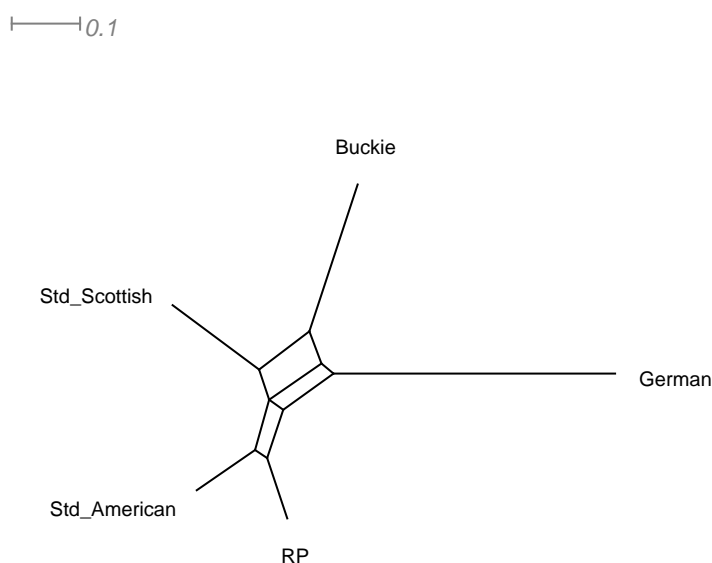
Figure 2.1: NeighborNet network for the Basic Edit Distance

first to apply Bayesian techniques from Evolutionary Biology to phonetic data. Heeringa (2004) used tree diagrams with clustering techniques and Multidimensional Scaling (Heeringa and Nerbonne, 2001; Heeringa, 2004). Principal Components Analysis is another option.

2.4.12 Application of NeighborNet in Small-scale study

The results of my small-scale comparison found that, as in previous studies (section 2.3.7 above), the different methods converged on a rather unified picture of the connections between the varieties. However, the NeighborNet diagrams also offer the opportunity to isolate a few of the areas in which the methods did not agree. I remind the reader that the subset of transcriptions used for this study come from McMahon et al. (2005-07) and are freely available on http://www.soundcomparisons.com (February 2011).

From figures 2.1 and 2.2, an extra reticulation distinguishes the basic Edit Distance from the Almeida & Braun method (and other feature methods). Warnow (p.c.) points out that NeighborNet is an implicit network, meaning that the reasons for the particular splits/branchings in the data are not overt (see also McMahon and McMahon, 2006). To discover them, one has to return to the raw data

⊢0.1
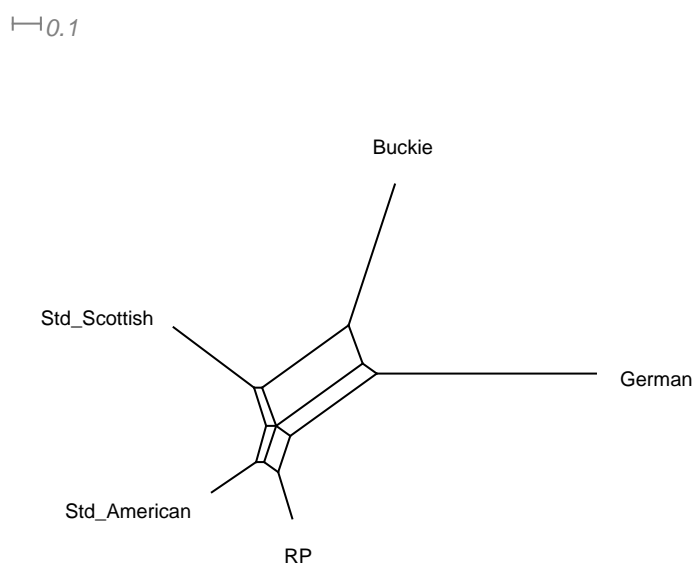
Figure 2.2: NeighborNet network for the Almeida & Braun method. The networks for the other feature methods were very similar to this.

and conduct Post-hoc analysis which is not ideal. By contrast, the Character-method Network explicitly demonstrates which Characters (words, sounds) account for the splits in the data. However, the difficulties I outlined in the previous section with getting phonetic data into a format acceptable for current Character methods place a barrier on appropriate Character analysis of the present phonetic data. Conducting post-hoc analysis on my data reveals that using the Basic Edit Distance, the pairs Std American English vs. Std Scottish English and RP vs. Scottish incur an almost identical overall distance score. However, in the feature methods, RP vs. Scottish receives a much larger distance than American vs. Scottish.

The key issue here is NeighborNet's propensity to make a network reticulation when it can (McMahon and McMahon, 2005, p.158).[16] This extremely important issue has not been thoroughly discussed. In order for NeighborNet to return a tree for phonetic data, the two varieties with the least phonetic distance would need to have an identical/almost identical distance score as each other to the next closest variety, and so on as the algorithm pairs up all the varieties in the set. Particularly with the feature methods, it would seem almost impossible, given

---

[16]Testing the robustness of NeighborNets with Bootstrapping is not straightforward, though McMahon et al. (2007, p.133) have done it.

3 varieties, for 2 of them to have almost equal distances to the third. It is thus inappropriate to view trees and networks as mutually exclusive, but rather as an issue of "resolution" (McMahon and McMahon, 2005). The issue that phonetic data will almost invariably produce a network rather than a tree structure shows the problem of using tools from other disciplines without being clear about the assumptions of these tools in relation to new data. It would be too easy to argue that the phonetic data support a network model rather than a tree model, especially as the debate between trees and networks in relation to language divergence is ongoing (Heggarty et al., 2010). I would thus also express more caution than McMahon et al. (2007, p.117), who have no problem using non-linguistic tools once their phonetic data are in numerical form.

2.4.13 *Feature methods behaving in a binary fashion*

There is value in using the degree of resolution provided by NeighborNet to examine how the different methods actually work. Returning to the discrepancy between the Basic Edit Distance and the feature methods outlined above, the well-known and salient rhoticity contrast among English varieties appears to be a strong factor. Std American and Std Scottish are both largely rhotic whereas RP is not.[17]

The rhoticity contrast in transcription largely involves the presence vs. absence of segments. This entails that the rhoticity contrast receives a large distance score in the feature methods as implemented in the present study (section 2.4.8 above). In this respect, the feature methods end up behaving quite similar to the basic Edit Distance in a rather binary fashion. This is especially true as other feature differences between the varieties (e.g. vowel differences) often incur quite small distances in the feature methods. These subtler differences do not make an impact alongside the large rhoticity scores. So essentially, the rhoticity contrast in the feature methods corresponds to a distance score of 1 in the basic Edit Distance, while the smaller vowel changes correspond to a score of 0. Although the feature methods are intrinsically more graded than the basic Edit Distance, in practice, they may echo the crudeness of the latter.

---

[17]There are also three examples ('cold', 'holy', 'one') in which American and Scottish both have mid-back vowels but RP has a central vowel in corresponding positions. Analysis of a larger data-set might show a pattern on a greater scale here. So rhoticity may not be the only explanation.

⊢ *10.0*
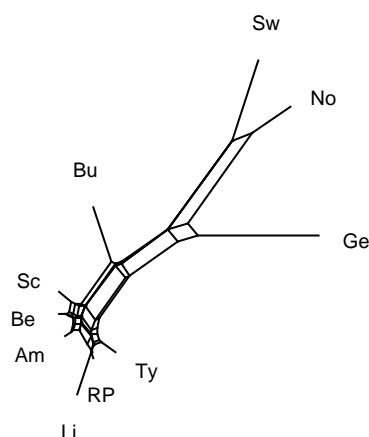
Figure 2.3: NeighborNet network for the 'Heeringa version' for the combined set of numerals 1-10

I wished to target this problem more specifically with a larger set of varieties (see section 2.4.1). In section 2.4 above, I outlined a different approach for scoring the presence/absence of a segment for each of the three feature systems and I compared these on this slightly larger set of varieties for the set of numerals 1-10.

The resulting NeighborNets are again broadly similar in topology (figures 2.3-2.5).[18] Post-hoc examination indicates that differences between the systems emerge in the numerals 'two', 'four', 'eight' and 'ten'. These differences again appear strongly related to how the different methods encode the costs for presence/absence of vowels and consonants respectively, rather than reflecting differences between the methods in gradations between segments. For example, in 'two', Liverpool's extra initial consonant [θ] leads it to occur opposite to Tyneside in a major split in the 'Heeringa version' and Connolly but on the same side in Almeida & Braun.[19] In 'four', the rhoticity contrast results in Buckie

---

[18]Abbreviations: Am (Std American English), Be (Belfast English), Bu (Buckie), Ge (Std High German), Li (Liverpool English), No (Norwegian (Stavanger variety)), RP (Received Pronunciation), Sc (Std Scottish English), Sw (Swedish (Stockholm variety)), Ty (Tyneside English (typical)).

[19]Liverpool: [t͡θʰɯː], Tyneside: [tʰʰɣ·u]. It is necessary to consult the NeighborNets for the individual numeral 'two' in appendix C to appreciate this point, not the NeighborNets for the combined set of numerals in figures 2.3-2.5.

⊢ *1.0*

Ty
Li
RP
Sc
Am
Be
Ge
Bu
No
Sw
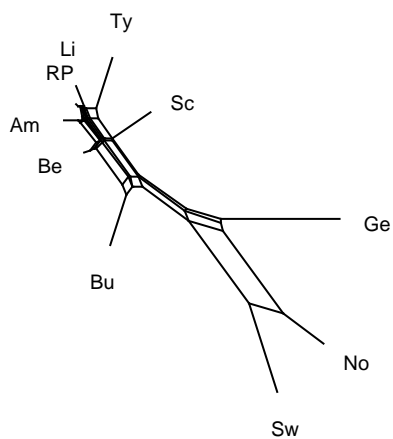
Figure 2.4: NeighborNet network for the Connolly method for the combined set of numerals 1-10

⊢──*10.0*

Sc
Be
Bu
Am
RP
Li
Ge
Ty
No
Sw
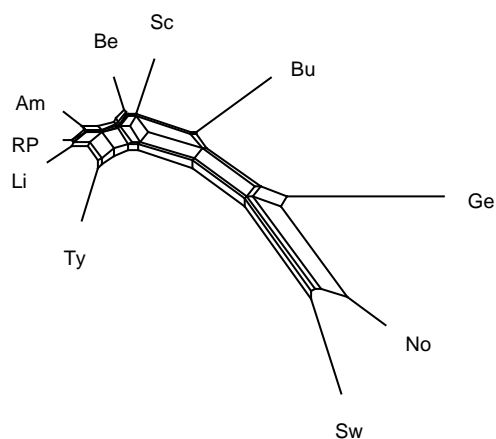
Figure 2.5: NeighborNet diagram for the Almeida & Braun method for the combined set of numerals 1-10

and German occurring on opposite sides of a major split in the Heeringa version and Connolly, but on the same side in Almeida & Braun (evident also in the networks for the combined set of numerals in figures 2.3-2.5). What this indicates is that the Heeringa version and Connolly systems give a higher cost to the presence/absence of a consonant than Almeida & Braun. In 'eight', Belfast and Scottish are on the same side of a split in the Heeringa version but on opposite sides the other two methods.[20] The Belfast pronunciation features a diphthong while the Scottish pronunciation features a short monophthong. This is a nice counterpart, indicating that the presence/absence of a vowel receives a higher cost in the Almeida & Braun system, than the Heeringa version, at least. The **sdists** function treats the diphthong/monophthong contrast as the presence vs. absence of the second vowel in the diphthong. Despite there being differences between the systems in how they score consonant/vowel differences, these differences do not come through, because they do not make an impact alongside the high scoring for the presence/absence of segments.

2.4.14 *Intermediate varieties*

NeighborNet reticulations can show intermediate varieties with links to other varieties in opposing directions. Buckie (see figures 2.1 and 2.2 above) has connections with German and with Std Scottish English. Although Buckie has retained certain archaic consonants now lost in the other English varieties but present in German (McMahon and McMahon, 2005, p.234ff.), it has undergone other sound changes which have affected English varieties and link its pronunciations with Std Scottish English in particular (including the Great Vowel Shift) (McMahon, 1991). Thus Buckie may be viewed as in an intermediate position, being pulled in some directions towards German and in others towards Std Scottish English. This also shows that the data is not entirely phonetically unified.

2.4.15 *Overall assessment*

This small-scale study comparing segmental methods of measuring phonetic distance has shown the following key points:

- The different methods performed very similarly overall on these data;

---

[20]Belfast: [iət], Scottish: [etʰ]. As with the numeral 'two', it is necessary to consult appendix C for the illustration of this point.

- Differences between the methods were evident only for high scoring contrasts like the presence vs. absence of segments;
- Articulatory gradations become lost.

Methods have a tendency to converge for a number of reasons. Minor phonetic differences encoded in the feature methods (e.g. small vowel height differences) do not make an impact on the overall score when alongside contrasts like rhoticity which incur far larger distance scores. Therefore, although feature methods differ in the scores they give to individual feature contrasts, these become levelled out. Another reason for convergence of different methods may be the process of aggregation and thus the expression of the distance between a pair of varieties as the aggregate distance score over the 30 words. Numerical values for different features could cancel each other out when combined together into an overall score. Particular feature contrasts may also need to be present in a large portion of the data in order to make an impact on the overall score. Using a much smaller set of features that are directly related to a particular hypothesis would be one important step in avoiding these problems (see section 2.3.6 above).

The basic Edit Distance takes no account of phonetic articulatory gradations, but yet gives a very similar picture to the feature methods of the relative distances between the varieties I examined. Given this situation, I argue that the outcome should not be considered an actual representation of *phonetic* distances between these varieties. The convergence of methods confirms that there is indeed a much closer connection of some kind between Std American English and RP English than between Std American English and German, for example, but it is not at all clear that this connection is a phonetic one.

Essentially the kind of overall 'distance' that even the feature methods largely produce is a count of the number of corresponding phonetic symbols which are not shared between two varieties in transcriptions of a given word. This does not constitute phonetic distance in the sense that it does not reflect degrees of articulatory proximity of individual segments.

2.4.16 *How to access 'true' phonetic distances*

How would I establish phonetic distances extracted from the quasi-distances being shown here? I would need to concentrate solely on those few areas of difference I have isolated in the performance of the methods, while at the same time rigorously remove the basic Edit Distance nature of the actual performance of

the feature methods. I could then set up a larger-scale study targeting these differences specifically. New approaches to scoring feature contrasts would also be necessary to ensure that the detail one encodes in the measure actually comes through. This would also ultimately involve getting away from orthographic-like segment/phones as units and making comparisons instead over longer and shorter spans than discrete segments. I return to this issue in chapter 5.

Admittedly, I may not have tested the different feature methods as thoroughly as would be needed. Mantel tests (see appendix B) on the distance matrices from the three feature methods showed much higher correlations between the vowels than between the consonants (though all correlations were significant at $p < 0.002$). This indicates that the differences between the methods were more in consonantal distances than vowel distances. Yet the majority of phonetic differences among English varieties at any rate lies among the vowels. Additionally, there may simply not be enough ways of articulatory feature systems being potentially different enough from each other for many phonetic differences that actually occur among varieties. The comparison of methods derived from clearly different standpoints (e.g. articulatory, acoustic, perceptual, or a combination of these (Mielke and Roy, 2008)) would provide more scope for seeing potentially different connections between the same set of varieties and scope for the assessment of these. Very little such comparison has yet taken place. Heeringa (2004) is an exception though his framework suffers from the problem arising from the presence vs. absence of segments that I have demonstrated in this chapter.

2.5 Acoustic segmental distances

Phonetic distance studies have been rooted with articulatory measurements and indeed Kessler (2005), in his review of such studies, hardly mentions the possibility of acoustically based studies. This is a further reason why my small-scale study involved the plain Edit Distance and articulatory feature methods only. Yet studies using acoustic measurements are increasing in two main fields: Dialectometry and Sociolinguistics. However, there has been little interaction between researchers from these two fields with the result that their methods and goals are rather different from each other. This section isolates some of the key points from segmental studies using acoustic distances. It then explores the similarities and differences between these studies and those using articulatory distances and between these acoustic segmental studies and my proposed approach to acoustic intonational measurements.

2.5.1   Discrete data

The first point to note is that even with acoustic distances, segmental studies often rely on discrete segmentation. This is true in the two main fields in which acoustic measurements are increasing: Quantitative Sociolinguistics and Dialectometry. Outside of these domains, Liljencrants and Lindblom (1972) measured acoustic distances between vowels in order to account for the distribution of vowel phonemes found in the world's languages. More recent work by Mielke, who is developing measures of articulatory, acoustic and perceptual distance in parallel, also relies on the presupposition that segments exist for his acoustic measurements (Mielke, 2005; Mielke and Roy, 2008).

In the sociolinguistic domain, Clopper and Paolillo (2006, p.449-50) made acoustic measurements from US English vowels in the isolated hVd context. Labov et al. (2006, p.37) also focused on analysis of pre-identified vowels in their goal of studying change quantitatively by age and gender. Boberg (2009) used a similar method to Labov et al. (2006) and analysed tokens of foreign /a/ in Canadian English.[21] In fact, the practice of using acoustic vowel analysis to show changes in progress has been well-established in Labov's work, e.g. Labov (1981) and vowel changes in Philadelphia English; Labov (2007) on the diffusion of changes, including splits and mergers and chain shifts. These kinds of changes have also been studied in British English, e.g. Kerswill (2003); Piercy (2010).

In Dialectometry, Heeringa (2004) and subsequent works, e.g. Gooskens et al. (2008) use phonetic transcriptions, as in the articulatory studies already discussed.[22] The difference is that instead of attaching numerical values to articulatory scales of place and manner of articulation, for example, they use spectrograms of individual vowels and consonants to derive distances between them. Heeringa (2004, p.81) cut samples of vowels and consonants from the tape The Sounds of the International Phonetic Alphabet by John Wells and Jill House. The vowels were produced in isolation and the consonants came from a small set of contexts and Heeringa acknowledges the simplistic nature of this. Heeringa and Joseph (2007) applied one of the methods used in Heeringa (2004) to the task of measuring divergence of Dutch dialects from reconstructed proto forms.

---

[21]This refers to words borrowed into English that contain one or more tokens of /a/, e.g. basmati.

[22]Heeringa (2004) included both articulatory and acoustic measurements.

Huckvale's work (Huckvale, 2004, 2007) used a broad phonological transcription of the data first, followed by phonetic segmentation with the HTK Hidden Markov Model toolkit. He also restricted his further analysis to vowel segments only. The most recent work comes from Ferragne and Pellegrino (2010), which builds on Huckvale's method but continues analysis of pre-segmented vowels. However, they did not begin with phonetic transcription. They isolated vowels through an automatic $f0$ detection procedure, though this resulted in part of the surrounding consonants being included in the vowel portion.

Heeringa et al. (2009) is one study which does not involve advance discretisation of the data through phonetic transcription. They made acoustic measurements between individual words in Norwegian, not between segmented individual vowels and consonants.

2.5.2 *Multiple tokens*

One key difference with articulatory studies of phonetic distance is that some of these acoustic studies use multiple renditions of vowel tokens from multiple speakers in their analysis, rather than accepting one transcribed form as representative for a given variety/language. This is the case for the studies falling into the sociolinguistic domain. Clopper and Paolillo (2006, p.449-50) used data from 24 males and 24 females and used five repetitions of each vowel per speaker. Labov et al. (2006, p.36-7) had circa three tokens of every vowel phoneme or allophone studied and had approximately 300 tokens from each speaker. In total, 134,000 vowel tokens from 439 US and Canadian English speakers were taken, putting this project on the grand scale. Boberg (2009) based his work on 1700 tokens of foreign (a). Piercy (2010) too used multiple speakers and multiple renditions of a given vowel token from a specific phonetic environment in her study of Dorset English.

Mielke's work (Mielke, 2005; Mielke and Roy, 2008) is based on recordings from four speakers (albeit phoneticians) and segments in three different contexts. Both Huckvale and Ferragne & Pellegrino's work use the Accents of the British Isles (ABI) corpus (D'arcy et al., 2004). There were approximately 10 male and 10 female speakers per variety (14 varieties in total) and five repetitions of each vowel token per speaker (Ferragne and Pellegrino, 2010). These repetitions were averaged.

The values of having multiple speakers include speakers behaving in different ways to expected ways. Ferragne and Pellegrino (2010) found half their Birmingham English speakers to have the FOOT/STRUT merger (Wells, 1982) (the expected pattern for this variety) and the other half not to have it. They also found some Glasgow speakers to be rhotic and others not. They argue this needs to be taken into account before getting averages. Isolated phonetic transcriptions of course cannot deal with this kind of interspeaker variability.

2.5.3 Measuring techniques

Sociolinguistically based studies generally rely on $f1$ and $f2$ measurements of vowel qualtiy, sometimes including duration. Clopper and Paolillo (2006, p.450) measured duration in milliseconds and took $f1$ and $f2$ measurements from the mid-point and 2/3 point of each vowel. Labov et al. (2006, p.36) used Linear Predictive Coding (LPC) to get a single $f1/f2$ measurement for the "central tendency of each [syllable] nucleus". They took an average of $f1$ and $f2$ values across the vowel. They did not get measurements of $f3, f0$, duration, intensity or bandwidth. They also performed normalisation to eliminate variability based on vocal tract length. Boberg (2009) conducted acoustic analysis using spectrograms and LPC. "Synchronous, nuclear measurements of F1 and F2 were taken at the maximal value of F1 or in the middle of a steady-state identified auditorily with the nucleus" (Boberg, 2009, p.364). He also performed a normalisation procedure on the vowel space. Piercy (2010, p.137) used LPC analysis in Praat to get $f1, f2$ and $f3$ values from vowel mid-points. She also took duration measurements and discusses speech rate normalisation procedures.

The dialectometrical studies use techniques explored by Heeringa (2004). Heeringa used three different methods of computing acoustic distances between vowels and consonants, two using spectrgrams and one using formant tracks. The spectrograms used the Barkfilter and Cochleagram representations (the former used in Heeringa and Joseph (2007); Gooskens et al. (2008)), both deemed to be closer to perception than the linear Hertz (Hz) scale because they use the Bark scale, which is like a logarithmic scale. Heeringa used Burg's algorithm to find the LPC coefficients for the formant track approach. He focused on $f1$ and $f2$ measurements converted from Hz into the Bark scale.

Unfortunately, Heeringa does not discuss how he dealt with consonants that do not have formant values and admits (p.84) that the bursts of voiced and voiceless

plosives will not match in his method. Heeringa et al. (2009) chose to measure formant tracks because in Heeringa and Braun (2003), these had higher correlations than Barkfilter spectrograms and Cochleagrams with perceptual distances and with phonetic transcriptions. Rather than measure formants in individual segments, Heeringa et al. unusually measured formants over each entire word. They measured the first three formant tracks. They also normalised for speech rate. Each word was broken down into frames and each frame contained three formant values. They compared corresponding frames from words in two varieties using the Manhattan distance. 'Silence' was needed for insertions/deletions of frames. Heeringa et al. also measured zero-crossing rates.[23] The difference between this work and other articulatory work by Heeringa and colleagues is that acoustic frames are compared, not phonetic segments.

In relation to articulatory distances (section 2.4.8) I discussed the problem of 'silence' for the Edit Distance and for the sdists procedure more generally. Heeringa treats 'silence' as a segment to be inserted, deleted or substituted in acoustic distances too. He cut a silent section from the same tape as he cut his vowel and consonant segments. He treated this as another segment and was thus able to calculate which vowels and consonants were closest to 'silence'. For example, in the Barkfilter representation, [ʔ] was closest to 'silence' and [a] was furthest from it. 'Silence' is a slightly less artificial concept in an acoustic context as opposed to an articulatory one i.e. frequencies and intensity values will simply be 0 (Heeringa, 2004, p.80). Nevertheless, it is still problematic to treat differences between the pronunciation of a given form in two varieties as involving 'silence'.

Although different segments have different numbers of spectra and formant bundles, Heeringa conducted a normalisation procedure (described on p.111) such that two segments being compared with each other had the same numbers of spectra/formant bundles. Like spectra, formant bundles just refer to single points in time. The former contain the intensities of the frequencies in that point in time. The latter contain formant values for that point in time (Heeringa, 2004, p.111). He then used the Euclidean Distance to compare pairs of spectra/formant bundles from two segments. He summed these and divided the resulting figure by the number of spectra/formant bundles. This gave a distance between two segments.

---

[23]The number of times the acoustic waveform crosses zero

The greatest distance between two segments was treated as 100% distance and other pairwise segment distances were expressed as percentages relative to this. The greatest distance was between [a] and 'silence' for both spectrogram representations and between [u] and [ʃ] for the formant tracks. Word and language variety distances can also be expressed as percentages. Interestingly, Heeringa does not mention whether he adopted a similar procedure to this important problem of deciding what constitutes 100% distance in relation to articulatory feature distances. Indeed a large problem with distance measures using articulatory segmental phonetic feature systems is that there is no clearly defined distance score that represents 100% distance on the scale. Further, there is very little acknowledgement or discussion of this problem in existing work (see also appendix B).

Heeringa, in his 2004 work, also calculated distances using acoustic representations of entire words. For this speech rate normalisation was needed (Heeringa, 2004, p.135). Here, Heeringa changes his approach from using just $f1$ and $f2$ to using $f3$ as well. The reasons for this are as follows:

- at each time sample, at least the first three formants could be found in every word;
- in whole word comparison, Heeringa did not make individual vowel and consonant comparisons against the IPA quadrilateral (p. 138).

However, Heeringa does not clearly explain his goal in deriving acoustic distances between individual segments, if his main interest was in fact to do so between entire words. It is also unclear exactly which elements were compared with which in relation to the entire words and whether these were valid comparisons. It appears that he used the same procedure here as in the later work (Heeringa et al., 2009) mentioned above.

Mielke used Dynamic Time Warping (DTW). Waveforms of each segment are broken down into overlapping slices of 15 ms yielding up to 200 slices per waveform. The spectra were then converted to Mel Frequency Cepstral Coefficients (MFCCs). This involved reducing the waveforms to 12 x 200 matrices. Then he calculated distances between pairs of matrices. DTW involves matching up intervals that are spectrally-similar. Their durations do not need to be the same. This yields a distance matrix with pairwise distances between each vowel and consonant.

Huckvale (2004, 2007) used LPC analysis to get formant locations for each half vowel. He used z-scores to normalise the formant frequencies for each speaker. He used average MFCCs for each half vowel. The ACCDIST measure first calculates pairwise vowel distances within an individual speaker using Euclidean Distances. It then gets correlations between the vowel distances of two speakers.

Ferragne and Pellegrino (2010) move away from the use of formant tracks pointing out their unreliability. They prefer methods drawn from speech technology which can capture more vowel information and enable large amounts of data to be processed. This reveals that the tension between linguistic faithfulness and computational ease is present in the study of acoustic segmental distances as well as the study of the plain Edit Distance and articulatory features. Like Mielke and Huckvale, they use Mel Frequency Cepstral Coefficients (MFCCs). They used 12 coefficients taken essentially from the vowel mid-point. Formant tracks can be wrong in the sense of the algorithm not finding the correct formant. MFCCs cannot be wrong in this way though they may be affected by background noise and other non-phonetic issues. MFCCs cannot be compared against existing phonetic values. Since their results fall into expected phonetic patterns, Ferragne & Pellegrino argue for the use of MFCCs over formants because MFCCs cannot be wrong. I disagree with this. Formants have the advantage of being phonetically transparent. Fruitful work could be done to improve the reliability of formant tracking and/or measuring how unreliable current methods of formant tracking actually are. The authors admit the further problem that it is unclear how the vowel dimensions of height and backness relate to their results using MFCCs.

Ferragne & Pellegrino's method was originally designed for automatic accent classification. They did not include duration in their measure. The reason why is that they did not know how to weight it against the 12 cepstral coefficients. They propose weighting features and including duration in future work. To get distances between pairs of vowels, they used the Manhattan Distance as this had the best results of accent classification over other metrics (e.g. Euclidean Distance) in their previous work. They made a matrix of pairwise distances for each speaker and then averaged then matrices of all speakers from a given variety.

2.5.4   *Analysing the results*

Clopper and Paolillo (2006) used Factor Analysis. This technique looks for correlations among several variables. It tries to find a set of uncorrelated factors

that account for the variation in the data. They performed two sets of analyses: the first on the vowel space, the second on talker variability. The first factor in the vowel space analysis was strongly related to $f1$ values, the second to $f2$ values. This leads the authors to assess the technique positvely for acoustic phonetic analysis (p.458). I agree that the transparency is an attribute in contrast to the ambiguous relationship between phonetics and MFCCs in Ferragne and Pellegrino (2010) (see section 2.5.3 above). In the talker variability analysis, formant values were also loaded on the first two factors but gender was also involved. There was an interaction between region and gender (p.454) such that males from the Northern US English region had higher $f1$ and $f2$ values than Southern males, a pattern not mirrored among the female speakers. Duration was heavily loaded on factor 3 and there was some link between vowel duration and region. The variables loaded on factor 4 were related to the $f2$ of high back vowels. This split the Midland, Southern and Western speakers from the New England and Northern speakers (p.457). Previous work already showed Southern and Western speakers with much fronted high back vowels. The 2006 work suggests Midland speakers may now be following suit. This illustrates that distances can help researchers to identify and analyse changes in progress. The variables loaded onto factor 5 were related to the Northern Cities shift which separates the Northern/Mid-Atlantic speakers from the Midland speakers.

Labov et al. (2006) used extensive $f1/f2$ plots. For example, to study the merger of /o/([ɑ]) and /oh/([ɔ]), they get distances between means on $f1$ and $f2$ plane of /o/ and /oh/ (p.62). Acoustic plots can also show the stages of the US Southern shift (p.73). They can also help to document the spread of changes (Labov, 2007), from different vowel locations for speakers of different ages (Apparent Time method).

Boberg (2009, p.365-7) focused on $f2$ values to decide the phonemic status of foreign (a) in Canadian English. Per speaker, he found the mean f2 values for /ae/([æ]) and /ah-o/(between [ɑ] and [ɔ]) and divided the distance between them arbitrarily into thirds. If a speaker's production of foreign (a) fell into the third nearest /ae/, it was classified as falling into the /ae/ phoneme. Likewise, if it fell into the third nearest /ah-o/, it was classified as belonging to the /ah-o/ phoneme. Tokens of foreign (a) falling into the middle third were considered extraphonemic. Using this approach, he found for example that US English speakers used far more /ah-o/ than /ae/ but it was the other way around for the Canadian English speakers. Further, he found extraphonemic tokens much

more among the Canadian English speakers and suggested on this basis that a new phoneme may develop. Again this shows how acoustic distances may be used to indicate processes of change.

Piercy (2010) used mean $f1$ and $f2$ values of BATH, TRAP (Wells, 1982) and other vowels in different segmental contexts. She is able to use this to chart four stages of change in which a one phoneme system with /a/ becomes a two phoneme system with /a/ and /ɑ:/, through lengthening and backing. These different stages can be seen in different speakers synchronically and is also a process known to have occurred in certain English varieties diachronically.

The dialectometrical studies generally look for correlations between their acoustic distances and perceptual scores. Gooskens et al. (2008, p.73-4) were interested in intelligibility of Danish speakers in relation to 17 Scandinavian varieties. As expected, they found a significantly stronger correlation between their intelligibility scores and consonant acoustic distances than between these scores and vowel acoustic distances. However, they only found consonant substitutions and vowel shortenings correlating significantly with the intelligibility scores. This highlights the discrete segmentation approach which pervades the dialectometrical studies, regardless of whether they use the plain Edit Distance, articulatory features or acoustic distances.

Heeringa et al. (2009) found the formant tracks with the Bark scale to have the highest correlations with perceptual distances for 15 varieties and the combined representation of formant tracks and zero crossing rates with the normalised Hertz frequencies to have the highest correlations for 11 varieties.[24] These results are for what they call semi-acoustic word measurements (because they rely on the phonetic transcriptions for speech rate normalisation). In relation to the fully acoustic measure, the combined representation had the highest correlations. The Bark scale had outperformed the Hz scale when frequencies were not normalised, but when they were, it was the other way around. Overall, the semi-acoustic measures had higher correlations with the perceptual scores than the fully acoustic measures. The semi-acoustic measures still make some use of phonetic transcriptions. Therefore, despite Heeringa et al.'s goal of moving away from transcriptions, they have only partially been able to do this.

The dialectometrical studies make much use of Multi-dimensional Scaling (MDS) and Cluster analysis. Non-linguistic issues like gender can emerge as strongly

---

[24]Normalisation was done to eliminate/reduce inter-speaker variability.

accounting for variability in the data (Heeringa et al., 2009). In Heeringa (2004, p.91ff), MDS plots (the first two dimensions) for the Barkfilter and Cochleagram corresponded well with the IPA vowel chart and consonant manner of articulation. Using the Mantel test, there were strong correlations between them (p.114). The formant tracks produced different results with a vowel triangle rather than a quadrilateral and no consonant distinction between plosives and fricatives. Heeringa's conclusion that formant tracks are not successful for consonants is entirely unsurprising given that many consonants do not have formants. Correlations with perceptual judgements on Norwegian data yielded the plain Edit distance with the highest correlation (0.67). The three acoustic methods with logarithmic distances were just behind at 0.66, while the Barkfilter nudged ahead of the others in Linear distances (Heeringa, 2004, p.186). As I discussed before(section 2.3.7), perception of varieties is extremely broad and includes several non-phonetic issues. Therefore, on a purely phonetic basis it is not appropriate to consider that the acoustic measurements were less successful than the plain Edit distance.

Heeringa and Joseph (2007) found found Frisian very similar to proto-Germanic in its consonant distances, while the Southeastern dialects had vowels close to proto-Germanic. I query how reliable the vowel reconstructions are, though, since there are more established procedures for the reconstruction of consonants.

Mielke (2005) also used MDS. He analysed the first dimension as corresponding to the distinction between "sibilant" and "sonorous", the second, to the distinction between "acute" and "grave", and the third, to the distinction between "low formant density" and "high formant density".

Ferragne and Pellegrino (2010) also use MDS and cluster analysis. First, they do this for the vowel systems of each individual variety. They advocate this method for examining phonemic mergers and splits. Afterwards, they do it for all varieties together, measuring the distances between the mean vowel spaces of different varieties. Huckvale was only interested in the latter, though generally their results concur. They assume the success of their method based on the fact that their results are broadly in line with traditional classifications of these English varieties. Huckvale does the same. They found some vowels to be very variable between accents and others less so. Their results for Southern British English indicate that the TRAP vowel has moved down and is merging with STRUT in some speakers (Wells, 1982). The MDS and cluster results concur in most but not all ways.

The danger of justifying a method against existing classifications of varieties is that it may not be an actual test. Warnow (p.c.) points out that even inferior methods can approximate the Indo-European tree (see section 2.3.7 above). Researchers do not ask enough what it would take for their method to produce a very different result to the traditional classification.

In this regard, Huckvale's further approach is more appropriate, I believe. To evaluate clustering, he examined whether all speakers from a given variety clustered together in their own sub-tree. He tested ACCDIST against a few other measures and found it to perform the best, though I observe that its best score is c.0.7. He did not find weighting made any difference. ACCDIST was most successful when using spectral envelope parameters, and Huckvale considers that this may be the best way to capture vowel quality.

2.5.5 Comparisons with articulatory measures and proposed intonational approach

One strong similarity between most of the acoustic studies reported here and articulatory studies is that both study vowels and consonants as discrete segments. Even in the case of the acoustic studies, this would involve first working with some kind of transcription (exceptions Huckvale (2004, 2007) and Heeringa et al. (2009)). Some of the acoustic studies compare multiple corresponding frames within individual segments, rather than just taking bundle of features per segment but this still maintains the same approach as the articulatory methods, only comparing smaller units. There is also the problem of treating 'silence' as a discrete segment-like unit in the acoustic distances of Heeringa (2004). DTW as used by Mielke (2005) is the equivalent of the Edit Distance for continuous data. The use of Manhattan Distance is similar between articulatory and acoustic works. Instead of comparing bundles of features, researchers compare bundles of spectra values, formant values or MFCCs. There may be a single bundle per segment or several. So even though the acoustic works are dealing with a continuous signal, they break it down into discrete units.

However, continuous acoustic measurements are able to show more graded relationships between segments than articulatory ones. For example, two vowels both classified with the features of [+high][-back][-round] would not have exactly the same $f1$ and $f2$ values. These more graded values make the acoustic distances more appropriate for studying changes in progress, which is exactly what the sociolinguistic studies discussed here have used them for.

The main difference between these approaches and the intonational approach I outline in chapter 4 is that in intonation, I do not work with transcribed data. I define a broader region (the Nuclear tone, see chapters 3, 4 and Appendix A: Glossary) than individual vowels and consonants and select two main acoustic points of interest within that region (L1 and H1, see chapter 4). It is these points that I assign a small set of measurements. In some respects, therefore, the intonational approach might appear to have some common threads with the acoustic comparisons of entire words in Heeringa et al. (2009). However, Heeringa et al. broke the entire words down into frames rather than selecting a few points of interest within the words. However, the acoustic segmental studies are similar to the intonational approach in that the acoustic values are taken from continuous scales in both cases, e.g. formant values, fundamental frequency (f0) values. It allows for a more refined expression of distance as the researcher is not simply forced to choose a discrete feature value to assign to a segment or intonational point. By contrast, articulatory methods limit the researcher to choosing from a very small number of values to represent vowel height, for example.

Acoustic studies generally use data from multiple speakers and analyse multiple tokens of the words/segments of interest. By contrast, articulatory studies tend to rely on a single transcribed pronunciation of each word/segment that is assumed to be representative of the entire variety. The acoustic studies have the advantage of being able to take into account that not all speakers of a given variety behave the same. They are also crucially able to take account of segmental contextual differences in vowel/consonant realisation. This information may be used to understand changes in progress (Piercy, 2010). In the intonational approach, I also use data from multiple speakers (albeit a very small number of speakers) and analyse several tokens per speaker. This also enables me to examine intonational variation according to prosodic context (the 'Short' and 'Long tails' in chapter 4) and I use this synchronic variation to suggest intonational change.

The questions of which elements to include in the measure and how they should be weighted are relevant to acoustic measures. The studies I have surveyed use a variety of different techniques, e.g. comparison of spectra, formant values, MFCCs. Duration was sometimes included, sometimes not. In my approach to intonational measurements in chapter 4, I focus on each acoustic measurement individually rather than attempt to combine them into a bundle of values. In chapter 5, I make preliminary attempts at combining the intonational

measurements into an overall score and reveal several problems with attempting to weight the elements against each other. The question of which elements to measure in intonation is not clear cut either. I motivate my choice of a few sparse measurements based on their grounding in the standard Autosegmental-Metrical (AM) theory. Chapter 4 reveals other potentially relevant measurements that for various reasons I had to leave aside for future work.

Heeringa's problematic use of 'silence' is extremely similar to his treatment of articulatory distances and the plain Edit distances. The issue of 'silence' is not relevant to the sociolinguistic research on specific individual vowels, nor to the other work discussed in this section (Mielke, 2005; Huckvale, 2007; Ferragne and Pellegrino, 2010). These latter works either also focus exclusively on vowels or only measured distances between individual segments. In this respect, these latter works have some similarity with my intonational approach because the specific points I measured were present in all the data I analysed, i.e. there was never an instance in which I compared the presence of an intonational point in one file with its absence (or 'silence') in another file.

In terms of evaluation, dialectometrical work uses both articulatory and acoustic measurements with perceptual scores. Correlations between intonational acoustic measurements and perceptual judgements is a topic for future work. It is not possible to study whether speakers of same dialect variety cluster together in the articulatory studies but this is possible with multiple speakers and acoustic data (Huckvale, 2007). One key goal of my intonational work presented in chapter 4 is to show a group of speakers (albeit a very small group) from the Glasgow English variety behaving in a way that is systematically different from another small group of Belfast English speakers. MDS and Cluster analysis are used in both articulatory and acoustic studies. These techniques are useful when the researcher has multiple varieties and/or multiple segments. Factor analysis has not been used with articulatory measurements, but could be used as an alternative to MDS. In my intonational work, I begin with just two English varieties and two intonational points (L1, H1). Therefore, these techniques would not have much to offer me at this stage. Phylogenetic networks have not been used with acoustic studies yet, though they have with articulatory ones. Ferragne and Pellegrino (2010) evaluate their results against existing classifications of varieties. Articulatory methods do this too. Agreement between their results and existing classifications is viewed as success of the measuring technique. In

chapter 4, by contrast, I use my intonational measurements to suggest that an existing grouping of English varieties (Urban North British) may not be as unified as once believed.

Acoustic methods can be used to assess degree of change from historical forms. They may also be used to show stages in phonetic and phonological change. Articulatory methods have not been used for either of these purposes but rather for the different historical goals of approximating cognate judgements and grouping varieties/languages together according to their historical relatedness. Chapter 3 outlines potential stages in intonational change and chapter 4 evaluates whether these changes have happened in the small amount of data from Glasgow and Belfast English that I examine. In intonation, I am not just interested in L1 or H1 comparisons for their own sake (unlike some of the acoustic segmental studies, e.g. (Mielke, 2005)), but for a broader understanding of differences between varieties and potential intonational change (somewhat akin to the acoustic segmental work of Ferragne and Pellegrino (2010) and Piercy (2010)).

2.6 What can be learned from segmental measures that could be applied to intonation?

The goal of this thesis is to extend measurements of phonetic distance into a new domain, that of intonation. I have just discussed acoustic segmental studies and have used them as a kind of bridge between the plain Edit distance/articulatory feature methods analysed earlier in this chapter and the intonational approach to follow in chapters 3 and 4. Yet do Edit Distances and their discrete letter-like operations in themselves have anything to do with intonation? What about the articulatory feature systems that I have outlined? At a surface level, these kinds of phonetic comparisons are very different from the phonetic analysis that has been done on intonation, as represented primarily by the acoustically gradient $f0$. However, the existing approaches to segmental phonetic distance provide some unexpectedly useful insights for beginning to approach the measurement of intonational distance.

Heggarty's Compatibility and Quantification problems would be as relevant to intonation as they are to segments. Heggarty used cognates to ensure that he could compare like against like. In intonation there are not obvious equivalents

to cognate forms. What are the other options? The unit I choose is termed the Nuclear tone and refers to the part of the intonation contour from the most prominent stressed syllable in the utterance until the end of that utterance.

In terms of quantification, I showed that gradable phonetic differences could become lost in a framework which decomposes segmental phonetics into discrete segment/phone-sized elements (section 2.4.15). The use of the Edit Distance algorithm and the **sdists** function more generally force approaches to the phenomenon of presence vs. absence of segments that are incompatible with actual processes of language change (section 2.4.8). This is also true of phenomena like metathesis (Heeringa, 2004, p.125-6). In chapter 3, I outline the possibility that intonational change may behave like metathesis in segments. If metathesis is vital to the account of segmental or intonational change, then it is crucial that our measure of distance can take this into account.

There is also increasing evidence that small phonetic differences may be important in intonation. For example, differences of less than 30 ms in the timing (Alignment) of peaks and valleys in the intonation contour may nonetheless distinguish varieties rather robustly (see Ladd et al. (2009) on the difference between RP and Std Scottish English intonation and Atterer and Ladd (2004) on Northern and Southern German). The importance of alignment differences in intonation must not be allowed to disappear in the scoring procedure for intonation. This also shows that choosing the few most relevant parameters in intonation and only measuring those may be the appropriate strategy here. My review of the existing literature on segmental phonetic distance combined with my own small-scale comparative study also concluded that including many features/parameters is problematic. Rather, I argue that we need to work with a few features at a time that are linguistically motivated and are directly related to very specific hypotheses.[25]

I have shown that the phonetic connections between varieties can be complex in relation to segments. There may be intermediate varieties showing conflicting links in both directions. Buckie was example of this, with its retentive connections with German and more progressive connections with Scottish English (section 2.4.14). Yet these intermediate varieties may be vital in helping to establish whether distances are likely to be indicative of an historical connection or not. This possibility may be relevant to intonation too.

---

[25]This may also address the problem of lack of independence of phonetic features, observed in relation to segments (see section 2.1 above).

2.7   Could intonation have any place here?

I now ask what it would take for intonation to have a place within the kind of framework that has been used with segmental measurements. The first major step is from distances derived from articulatory feature systems to distances based on acoustic measurements. These acoustic measurements would also need a different approach to the existing acoustic segmental measures discussed in section 2.5 above.

There is some evidence to suggest that certain intonational phenomena behave categorically in certain contexts (Pierrehumbert and Steele, 1989), while others are intrinsically gradient (Ladd and Morton, 1997; Atterer and Ladd, 2004). Some of the most recent work on intonational parameters comparing varieties indicates that there may not be discrete boundaries representing categorical differences between varieties (Mücke et al., 2009). Though I have outlined problems with an overall aggregate score of distance (section 2.4.11), it is certainly a valid question how intonational parameters could be combined into an overall score. Could categorical and gradient be combined into a single score? In segments, bundles of features have been regularly combined but it is much less clear how this should be done with acoustic intonational parameters. A small set of features like segmental phonetic features have made their way into certain intonational analyses (Ladd, 1983; Gussenhoven, 1984). However, current phonetic analysis relies more on actual exact measurements along the time and frequency scales. Since acoustic intonational parameters are measured on these different scales, how would I reconcile measurements from different scales? If I did reconcile the different measurements, I could use the Manhattan distance as I used for segments (see section 2.4.4 above and chapter 5).

In section 2.4.3 above, I identified a problem with most feature approaches. They do not justify the numerical values attached to individual features. Choosing how intonational parameters should be weighted against each other would be a great challenge. For hypotheses about historical change in intonation, it would be necessary to have strong hypotheses about the relative historical stability of intonational parameters. At present, there are not strong indications of such historical stability.

In the Autosegmental-Metrical (AM) theory, which has become the dominant theory of intonation, the inventory of elements from which intonation contours are composed is more minimalistic than in segments. Combined with the lack

of an established framework for studying intonational change and the lack of historical evidence in intonation, it would thus be much harder to distinguish chance similarities from historical connections. There is an unexpected parallel here with Longobardi and Guardiano (2009), who have begun to use phylogenetic techniques with syntactic data. They assess syntax as having lesser variability than the lexicon and "hence similarities are less probative" (p.1684).

2.7.1 Potential benefits

A crucial question is whether a distance-based approach would actually have anything to offer intonational phonetics. I believe that the largest potential benefit could be in exploring and testing new hypotheses (developed theoretically) about the potential historical stability of intonational parameters. Existing methods used with segments allow the inclusion of only the most stable elements (either the meanings argued to be most stable and/or the features). It is equally possible though to compare elements argued to be more or less stable and examine the resulting data patterns (McMahon and McMahon, 2005). An argument against this in intonation would be that the fewer intonational parameters (compared with the number of segments and features) would be transparent enough in their own right. Chapter 5 illustrates some preliminary attempts at weighting the intonational parameters of alignment and scaling and deriving an overall score of intonational distance.

If I did derive a combined intonational distance measure capable of expressing distance between varieties, then it could be the input to a program like NeighborNet in exactly the same way as segmental distance measures. This does not entail that the NeighborNet output would necessarily be meaningful. The results would need to be carefully scrutinised alongside measurements from the individual parameters and accompanying theoretical arguments about potential historical change and connections between the varieties. In section 2.4.11 above, I showed how the nature of the data from segmental phonetic measures biased the production of a reticulate network diagram rather than a tree diagram with NeighborNet. Therefore, I caution against any rash uses of NeighborNet to support arguments in fields outside the biological field for which it was developed.

This represents a very slender outline of the potential for development of measures of intonational distance, given the background of the existing segmental approaches. There are clearly large gaps between the two. There is the difference

between discrete and continuous measurements. Most existing articulatory segmental measures are carried out on phonetically transcribed data, with one transcription per word per language. Though acoustic segmental studies typically involve multiple speakers, most too rely on discrete transcriptions. Intonational measurements by contrast are usually continuous acoustic measurements on a large number of utterances from each variety/language and generally use more than a single speaker. Whether intonational data can be appropriately discretised is ambiguous. The sharp differences between intonation and approaches to segmental distance do not mean that measures of intonational distance are beyond reach though.

2.8 Why intonation has been left aside by historical linguists and dialectometrists

Intonation has been a very marginal interest within Historical linguistics. Exceptions include Colantoni and Gurlekian (2004), Gussenhoven (2004, p.225), Elordieta and Calleja (2005) and Hualde (2007). No doubt the primary reason for this is the lack of historical records and established historically-oriented frameworks in intonation. This stands in contrast to segmental phonetics/phonology and to lexical tone. The fact that intonation is not preserved in writing is one factor but not the only one, I argue. The study of historical sound change has the benefit of the very long-established traditions of the linguistic Comparative Method and Internal Reconstruction (Joseph and Janda, 2003). Studies of language contact too have generally left intonation to the sidelines until recently (Queen, 2001; Hualde, 2007) and it is unclear whether contact would work in similar ways in intonation and segments (McMahon, 2004). What segmental phonetics/phonology and lexical tone have in common is that they have the word/morpheme as a unit. As I demonstrated in section 2.3.1 above, this has provided a ready unit in the form of the cognate for the implementation of quantitative methods. It directly addresses the Compatibility problem of Heggarty (2000a). In intonation, it is not as clear what we should be comparing against what. In section 2.7 above, I indicated that I shall consider the 'Nuclear tone' to be the 'unit', and I justify this at greater length in chapter 3.

Within Historical linguistics, certain phonetic distance measures have the explicit goals of the identification of ancestral relationships between languages and the identification of cognate forms (section 2.3.7 above). The small time

depth at which intonation may be studied prevents any relevance of intonational comparison to the former goal. Likewise, the cognate being defined at the word/morpheme level makes no reference to intonation. Heggarty's method relies on the existence of cognates and any comparison beyond this is a problem for him. Heggarty (2000a, p.114) only includes a very brief proposal of potential parameters that could be used in measuring intonational similarity/distance.

Like Historical linguistics, Dialectometry has been rooted in the study of segmental phenomena. In section 2.1 above, I showed how both fields have taken lexical comparison involving word/meaning lists as a model for how to approach phonetic distance. Separately, prior to the development of quantitative methods of lexical and phonetic distance, there was the tradition of isogloss drawing (Chambers and Trudgill, 1980). This tradition did not concern itself with intonation either.

Dialectometrists have the goal of matching phonetic distances with perceptual judgements either of experts or of naïve listeners. This does not preclude intonation. Intonational characteristics are salient for listeners in identifying dialects (Peters et al., 2002, 2003), although that is a somewhat different goal to assessments of distance between varieties. Gooskens (2005) examined the role of intonation in dialect identification in Norwegian. She included this alongside measurements of segmental distance between dialects using the Edit Distance. However, she did not attempt measurements of intonational distance. Dialectometrists do not have a background in the prominent frameworks for intonational analysis. Thus they would probably struggle to adapt their existing measures based on articulatory phonetic features. The phonetic realisation of intonation is not so amenable to an IPA-based discrete, feature representation. Even the acoustically based measures discussed in section 2.5 above would not easily be adapted to be linguistically-meaningful in intonation without thorough knowledge of intonational frameworks. Additionally, Dialectometry takes no account of interspeaker variability, which I highlighted was important in intonation (section 2.7 above).

Since intonation can add little or nothing to the understanding of sound change at deep time depths, it is understandable that studying intonational change would not have been an appropriate goal of those historical linguists who have used phonetic distance measures. The primary reasons why dialectometrists have also been reluctant to embark on large-scale intonational measurements at present are:

- the prevailing framework of isolated word comparison and discrete transcriptions;
- dialectometrists not having a background in intonation or other aspects of prosody.

2.9   Approaching Chapter 3

This chapter has shown that existing measures of phonetic distance involve segments and not intonation. I believe strongly that intonation deserves a place within the study of measurements of phonetic distance. Acoustic intonational parameters offer plausible possibilities for assessing distance between varieties of a language. Historical questions are appropriate in intonation. The study of intonational variation among varieties is increasing. Thus there is potential for considerable headway in using intonational distance measurements alongside hypotheses about intonational change. Chapter 3 sets out the theoretical context for measuring intonational distance and evaluating intonational change.

CHAPTER 3

Potential intonational change

> I have only observed that he generally spake the last syllable in a sentence nearly a third above the last but one (said in 1789 of a Tyneside boy. Source: Ellison Collection, Gateshead Library, quoted by Hughes (1956, p.36) and cited in Cruttenden (1995, p.160)).

3.1 Overview of chapter

Chapter 2 concluded with a discussion of why historical linguists and dialectometrists have not examined intonation. Chapter 3 first asks why phonetic distance measurements have not been of interest to intonation specialists. Next, I introduce my central interest in intonation, the phenomenon of rises on statements in Belfast and Glasgow English. In keeping with the Historical linguistics dominance in segmental phonetic distance measures, I have chosen a quasi-historical context for my approach to intonational measurements. Statement rises in Belfast and Glasgow have been of interest since Jarman and Cruttenden (1976), yet their origin is uncertain (Cruttenden, 1995, 1997). Therefore, they provide an ideal case study for examining intonational distance and change. I set out my initial hypothesis that the statement rises have changed over time from statement falls. I then put forward a theoretical account of the historical origin of these rises consistent with this hypothesis. An experimental evaluation of this theoretical account involving intonational phonetic distance measurements follows in chapter 4.

3.2 Why phonetic studies on intonation have not focused on measurements of phonetic distance

Just as dialectometrists have largely eschewed intonation, intonation specialists have avoided the kind of phonetic quantifications that have arisen in Dialectometry. I see two primary reasons why this is the case. The first reason, as explained in chapter 2, is that intonation has not had a place within the historical linguistic and dialectometrical frameworks from which many measures of segmental phonetic distance have sprung. The second is that the key goals of major works on intonation in the last 50 years or so have been in other areas than the numerical expression of degrees of relative distance either between intonational units themselves or between varieties/languages.

3.2.1 Problems of cross-linguistic comparison

Though phonetic analysis of the intonation of varieties of a single language is increasing, direct cross-linguistic comparisons are not commonplace. At the heart of this is the difficulty of deciding what should count as instances of the 'same' contour/element across languages. Although the ToBI (Tones and Break Indices) system of intonational transcription was originally designed to be like the IPA (Silverman et al., 1992),[1] it became clear from later work that ToBI is very much language and variety specific, must be adapted to fit the intonational characteristics of different languages and varieties and is actually a *phonological* transcription system (Beckman and Ayers Elam, 1997; Mayo et al., 1997; Veilleux et al., 2004). Indeed Mayo et al. (1997), Grabe (2002) and Ladd (2008, p.128ff.) show that the Belfast and Glasgow rises (of central interest in this thesis) cause problems for the standard ToBI system.

Hirst and di Cristo (1998) is the only attempt to create an IPA-like system specifically for cross-linguistic comparison in intonation (INTSINT). These researchers wished to isolate the "different dimensions" in which languages differed intonationally. However, Hirst & di Cristo's work has remained on the fringes of mainstream intonational research. The primary reason I can identify for this is that

---

[1]"Unfortunately there is no single standard for prosodic transcription that is analogous to IPA for phonetic segments. To meet this need, a group of researchers with expertise in a variety of approaches to prosodic analysis and speech technology have developed ToBI: an agreed transcription system which builds on much recent progress in prosodic modelling" (Silverman et al., 1992, p.867).

Hirst & di Cristo have been far too all-encompassing in their attempt, including metre, prominence and different elements of pitch from the global overall contour shape to local comparisons between the height of adjacent $f0$ peaks or valleys. With this degree of breadth, it is difficult to ascertain which aspects of intonation are worthy of cross-linguistic comparison and which are representative of the range of variability that may be expected within a single language.

3.2.2 *British Tradition, American Structuralist School and Dutch Tradition*

The primary schools of intonation analysis have concerned themselves with the kind of meaningful quasi-phonological units into which the intonation contour should be decomposed. In the British tradition of intonation analysis, there has been a prominent focus on the teaching of intonation contours to non-native English speakers (O'Connor and Arnold, 1961). The American Structuralist school (Trager and Smith, 1951) has attempted to find intonational equivalents to the concept of the phoneme in segmental phonology. Researchers within these traditions have also generally had little access to tools which enable systematic acoustic analysis of $f0$ contours (see also Ladd, 2008, p. 11). Quantifying phonetic distances was not a goal and, in any case, attempts would necessarily have been impressionistic.

By contrast, the Dutch tradition has used instrumental techniques from the outset ('t Hart et al., 1990). Quantifications of phonetic distance have indeed been part of the analysis here, but with the goal of assessing what kinds of modifications there could be to intonational contours such that listeners would still perceive the contours as 'the same'.

3.2.3 *Autosegmental Metrical (AM) theory*

In the last 30 years, Autosegmental-Metrical (AM) theory has become dominant in intonation analysis (Liberman, 1979; Pierrehumbert, 1980; Ladd, 2008).[2] AM theory too has tried to isolate the phonological building blocks of intonation. It is notable for its minimal approach of only distinguishing between High (H) and Low (L) tones. It decomposes the intonational contour into Pitch accents (associated with prominent stressed syllables) and Edge tones (associated with prosodic

---

[2]The work of Bruce (1977) on the Swedish Lexical accent was deeply influential on the development of an AM treatment of intonation in English (see Gussenhoven (2004, p.209) and also Horne (2000)).

boundaries such as the IP). Several studies have now examined the phonetic realisation of these tones in various languages, but often with the goal of trying to understand the phonological categories in intonation (Pierrehumbert and Steele, 1989; Arvaniti et al., 2000, 2006; Grice et al., 2000), which phonological contrasts are present in the data (Prieto et al., 2005), or with the goal of identifying the phonetic correlate of a phonological tone (Ladd and Schepman, 2003). Therefore, the goal has not been on phonetic distances themselves.

3.2.4 AM phonetic analysis

There are two primary phonetic parameters in AM phonetic analysis:

- Alignment
- Scaling.

Alignment refers to the precise timing of H and L targets with respect to segmental and syllabic landmarks, e.g. measuring how far from the beginning of a stressed vowel onset a local $f0$ maximum occurs. Scaling refers to the relative height of H and L target points in the individual speaker's pitch range at that point in the utterance (Ladd, 2008).

Figure 3.1 is intended to clarify what alignment and scaling are and how they are measured.

Acoustic phonetic data has offered researchers the opportunity to argue for a representation of intonational tones as H and L targets (local $f0$ maxima and minima), in contrast to the overall shapes of 'rises' and 'falls' which were the building blocks of the British tradition of intonational analysis (Arvaniti et al., 1998). Overall contour shapes result from the alignment of the individual L and H targets with the segmental material (Arvaniti et al., 2006; Ladd, 2008). Another strong interest has been in how the phonetic realisation of these tones (particularly H tones) changes in different prosodic, segmental, and other contexts (Steele, 1986; Silverman and Pierrehumbert, 1990; Bruce, 1990; Prieto et al., 1995; Rietveld and Gussenhoven, 1995; House et al., 2000; Ladd et al., 1999, 2000; Knight and Nolan, 2006; Schepman et al., 2006; Arvaniti and Ladd, 2009; Ladd et al., 2009; Mücke et al., 2009).

One of the most robust findings across studies has been the alignment difference between H targets associated with the primary stressed syllable (Nuclear)
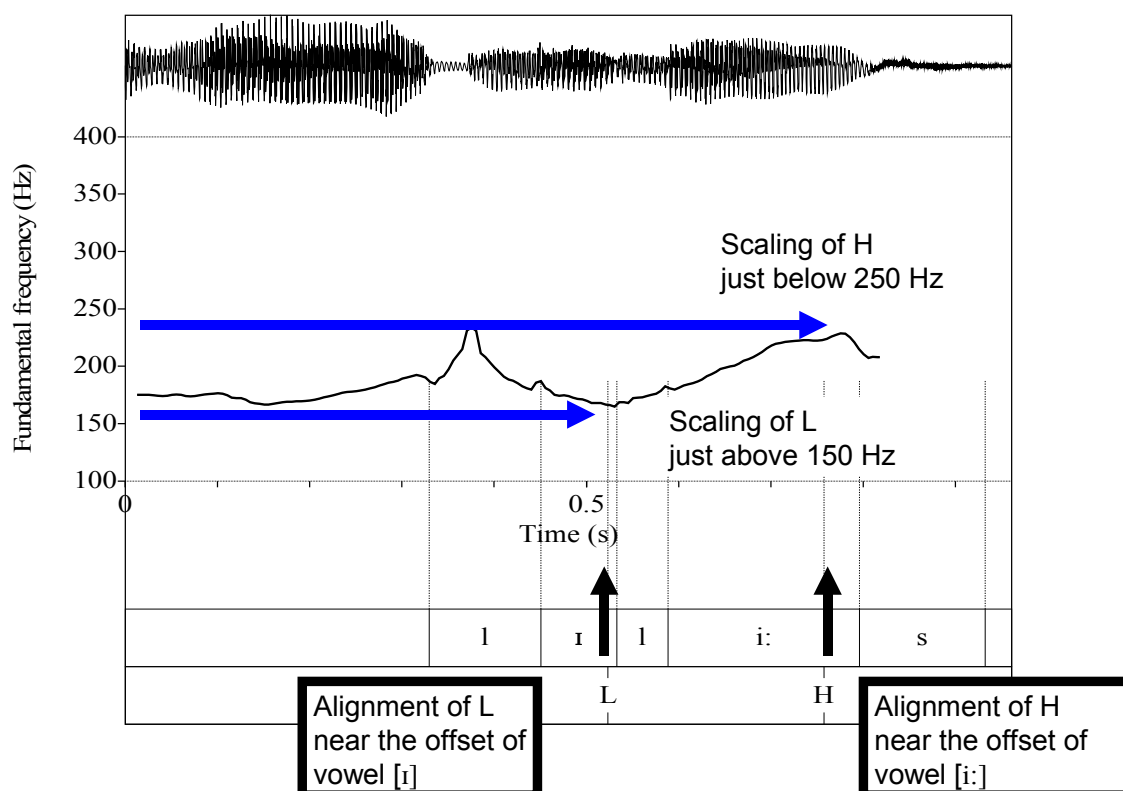
Figure 3.1: Waveform and *f*0 trace of the word 'lilies' in nuclear position spoken by female Belfast English speaker RB. This is an extract from Experiment II in chapter 4. The scaling values here have not been normalised by the speaker's pitch range.

and H targets associated with preceding stressed syllables (prenuclear). The difference is that the alignment of H is consistently earlier in nuclear syllables than the alignment of H in prenuclear syllables. One attractive explanation is tonal crowding. When two intonational tones come too close to one another, there may be phonetic effects (Silverman and Pierrehumbert, 1990; Arvaniti et al., 1998, 2006; Mücke and Hermes, 2007; Arvaniti and Ladd, 2009; Ladd et al., 2009; Mücke and Hermes, 2007).[3] In this instance, the nuclear H may be adjacent to an edge tone leading to crowding (Ladd, 2008).

Often the study of prosodic/segmentally conditioned variation has overlapped with the previously mentioned goal of trying to match phonetic targets to phonological tones (Arvaniti et al., 2000, 2006). An increase in the number of preceding (Nolan and Farrar, 1999; Dalton, 2007) and following unstressed syllables (Steele,

---

[3]Again Bruce's work (Bruce, 1977, 2007) on the Lexical accent in Swedish has been particularly influential here.

1986; Silverman and Pierrehumbert, 1990; Dalton, 2007) often results in respective leftward and rightward movements of H. The rightward movement of H under the influence of following unstressed syllables is an important point, to which I shall return below (section 3.4.2). The alignment of H targets can change alongside changes in phonetic vowel length (Steele, 1986; Silverman and Pierrehumbert, 1990; Knight and Nolan, 2006; Mücke and Hermes, 2007) and phonological vowel length (Ladd et al., 2000; Schepman et al., 2006; Ladd et al., 2009). Consonantal effects have also been well-studied (Arvaniti et al., 1998; Mücke et al., 2009).

Less work has examined the intonational realisation of the fundamental sentence type distinction of statement vs. question, though Makarova (2007) and D'Imperio et al. (2007) are exceptions. A repeated finding though is that questions have a higher pitch and/or wider pitch range than corresponding statements (e.g. Yuan, 2006). There may also be speech rate differences between statements and questions (van Heuven and van Zanten, 2005).

From this overview, it is clear that *degrees* of difference in phonetic realisation have not been a focus of existing work on intonational phonetics. To date, phonetic analyses of intonation have been particularly concerned with the mapping between the phonetic realisation of H and L targets in the intonation contour and the phonological H and L tones which make up the 'units' in AM theory. Braun (2006) has measured the realisation of intonation along several different parameters, but more with the goal of isolating the most relevant cues to phenomena like focus marking. From this broad overview, it is also evident that studying intonational change has not been a priority in any of the major schools of intonation. However, it is worth probing whether the phonetic changes that are well-documented in different prosodic and segmental contexts (e.g. tonal crowding) could also occur in a diachronic context too.

3.2.5   *Segmental Anchoring Hypothesis*

A dominant hypothesis in AM analyses of intonation has been the Segmental Anchoring Hypothesis (Ladd et al., 1999, 2000; Ladd, 2006; Dilley et al., 2005), which argues that H and L target points are independent of each other and are each tightly coordinated with the segmental material. This hypothesis explicitly argues against the notion that two tones are connected by a constant slope or duration (Arvaniti et al., 1998). This is the fundamental point. Contra Prieto and

Torreira (2007), the hypothesis does not predict that the alignment of H and L target points should be with syllable or vowel edges. However, results even from the above mentioned studies do not always clearly support this hypothesis (e.g. Ladd et al., 2009, p.151). The co-varying of scaling of L and a following H in Ladd and Schepman (2003) would also not fully fit with an argument that L and H are independent of each other. Separately, other studies have argued for a much weaker version of this hypothesis (e.g. Welby and Loevenbruck, 2006). The most important new direction in the testing of the Segmental Anchoring hypothesis is the finding that H and L targets may be anchored to *articulatory* events rather than *acoustic* ones (D'Imperio et al., 2007; Mücke and Hermes, 2007).

3.2.6 *Diachronic change in intonation, lexical tone and lexical accent*

Very few researchers have used the subtle phonetic measurements involved in the parameters of alignment and scaling to explore potential change over time in intonation. Considerably more work in this area has been done on Scandinavian languages with Lexical accent (Swedish, Norwegian) (Engstrand and Nyström, 2002; Bye, 2004, 2008; Hognestad, 2006). There has also been work on diachronic tonal height changes in lexical tone in Chinese languages (Chen, 2000).

The diachronic research on the Scandinavian languages is a direct follow-on from synchronic research comparing the alignment of Lexical accents I and II in varieties of these languages (Bruce, 1977; Kristoffersen, 2006, 2007). Bruce (1977) did extremely important influential work on Stockholm Swedish, showing that the fall after the H peak in accent I occurred in the consonant onset of the stressed syllable, whereas the fall in accent II occurred later, in the stressed vowel. The alignment difference between H in accent I and II is the fundamental distinction, though there are other phonetic differences (e.g. a sharper fall in accent I than accent II), which have received much less attention.

Segerup and Nolan (2006) go against the view that the Swedish accent distinction is a difference solely of the alignment of H. In Gothenburg Swedish, they found the H of accent II to be consistently higher than the H of accent I. Perceptual tests mirrored this. To account for their findings that the accent distinction could be marked either by alignment or by the scaling of H, Segerup & Nolan invoked 'Cue Trading'. This gives the speaker/listener the opportunity to choose from two or more strategies for signalling the accent distinction. Cue trading considers the phonetic dimensions to be independent (Repp, 1982), but alignment

and scaling are not always independent of each other (see section 3.2.9 below). Further, Segerup & Nolan's results do not actually show evidence that scaling is clearly taking the place of alignment nor the other way around. Nevertheless, the findings of other differences between the Swedish accents than the key alignment difference is a reminder that an alignment difference on its own may also be unlikely in intonation.

3.2.7 Stability and variability in intonation

The dearth of historically oriented phonetic analyses of intonation mean that evidence on diachronic stability/variability of intonational tones and key parameters is slight. However, there are reports (albeit sometimes conflicting ones) on the stability and variability of these elements in a strictly synchronic context. Many researchers have proposed that H tones may be more variable in their phonetic realisation than L tones (Silverman and Pierrehumbert, 1990; Caspers and van Heuven, 1993; Prieto et al., 1995; Ladd et al., 2000; Ladd and Schepman, 2003; Dilley et al., 2005; Prieto and Torreira, 2007). However, this may be a consequence partly of more research having been done on H tones.

In relation to the parameters themselves, Arvaniti and Ladd (2009) present alignment as more variable than scaling in relation to the specific phenomenon of tonal crowding. However, Xu (2005) argued that scaling was more variable than alignment. The languages studied in these two respective studies were completely different, Greek in the former, Mandarin Chinese in the latter.

There is an added dimension that aspects of scaling are paralinguistic. For example, pitch raising may be used for marginally linguistic purposes such as the expression of emphasis and signalling the emotional state of speaker. Thus perhaps scaling is somewhat outside the remit of *linguistic* change. Not many recent studies have examined how stable the scaling of peaks is (Ladd, 2008), though Connell and Ladd (1990) found that Yoruba speakers' scaling of peaks was under very careful control. One very familiar finding though is that in questions, H targets are higher scaled than in corresponding statements, or have a wider pitch range than statements (Haan, 2002; Yuan et al., 2002; Gussenhoven, 2004). There is much less consistency in the behaviour of the scaling of L, suggesting that the scaling of L may be more variable than scaling of H. Further support for the greater variability of L scaling comes from Arvaniti and Ladd (2009) and

earlier work by Arvaniti and colleagues, who found that the scaling of H targets was not as strongly affected by tonal crowding as the scaling of L targets. However, in the absence of tonal crowding, L targets can be consistently scaled (Ladd and Schepman, 2003; Dilley et al., 2005). The apparent greater stability of H targets may indicate that the realisation of H targets is more important than the realisation of L targets (Grice, 1995; Arvaniti et al., 2000).

There is an unresolved issue here. What counts as an equivalent amount of change in relation to alignment and scaling, since both are measured on different scales? This has strong implications for claims that alignment is more variable in the sense of exhibiting greater changes than scaling (a claim made by Arvaniti et al. (2006)). The most obvious strategy would be to invoke perceptual thresholds for both alignment and scaling to arrange some kind of equivalence between them but this has not been done (see chapter 5 though).

What I conclude from this exploration of the relative stability of alignment and scaling is that it is not possible to draw a simplistic distinction between stable scaling and variable alignment.

3.2.8 Cross-variety differences in intonation

Researchers have shown a burgeoning interest in cross-variety differences in the alignment of H targets in particular (Bruce and Gårding, 1978; Atterer and Ladd, 2004; Dalton and Ní Chasaide, 2005; Elordieta and Calleja, 2005; Arvaniti and Garding, 2007; Dalton, 2007; Ladd et al., 2009; Mücke et al., 2009). Ladd et al. (2009) found, for example, that Standard Scottish English speakers have peaks (H) later timed in the accented syllable than Southern British English speakers. However, these variety differences may be gradient rather than categorical (Atterer and Ladd, 2004; Mücke et al., 2009). The goal of these researchers has been the understanding of how the range of variability among varieties can be integrated into the topic of phonological categorisation in intonation. There have been no explicit attempts yet to measure intonational distance between two varieties. This thesis attempts to address this in the context of English varieties, particularly Belfast and Glasgow English.

3.2.9 Lack of independence

A potential problem with using alignment and scaling in a measure of intonational distance is that they are often not independent of each other. Pitch range[4] (the difference between the scaling of L and H) can affect alignment (Knight and Nolan, 2006). One recurring finding is that H peaks tend to be later aligned alongside an increase in the height of H or a widening of the range between L and H (Ladd and Morton, 1997; Knight and Nolan, 2006). In addition, raising the height of H can function in the same way as as timing the H later (House, 2003). Makarova (2007) actually found *earlier* alignment of H working together with high pitch in her perceptual experiments. Alignment and scaling do not always affect each other though (see Knight and Nolan, 2006, p. 31 and Arvaniti et al 2006).

3.2.10 Existing attempts at measuring intonational distance

Surprisingly, there are very few measures of intonational distance between 'real' and synthesised contours in Speech synthesis (an exception is Clark and Dusterhoff (1999)). Actually, intonation has an ambiguous place in current synthesis models because adding too much phonetic detail to the model can actually reduce the naturalness of the resulting synthesised speech rather than improve it. Taylor (1992, 2000) stands out for phonetic modelling of intonation in a speech synthesis context. His work includes formulae to quantify the amount of rise and fall around stressed syllables and intonational boundaries. Taylor's work did not include any cross-variety study, but in principle, his techniques could be applied to comparisons between varieties (see chapter 5). Many of the phonetic studies on alignment were done with a goal of producing Text-to-Speech systems for that language/variety (e.g. Silverman and Pierrehumbert, 1990; Prieto et al., 1995; Rietveld and Gussenhoven, 1995; Kochanski and Shih, 2003; Dalton and Ní Chasaide, 2005; Arvaniti et al., 2006). The findings on the alignment of the H and L target points could then be expressed as rules in speech synthesis systems.

Though not with the goal of measuring distance, House and Wichmann (1996) and House et al. (2000), drawing on Rietveld and Gussenhoven (1995), did derive a simple scoring system for predicting the alignment of H targets within stressed syllables. This was based on the segmental composition of the syllable and the effects that onsets and codas have on alignment. House and Wichmann

---

[4]Also termed pitch span or $f0$ excursion, though these terms are used in a variety of ways.

(1996, p.114) admit that their scoring system is a "rather crude unitary measure" and they do not justify the three broad categories into which they group their scores. Nevertheless, a scoring system along the lines of the one they outline might be appropriate for measuring distance between two intonation contours or between varieties of a language. A few studies have attempted to measure distance in intonation, but these remain isolated pockets in the literature. Connolly (1997) tentatively outlined how the Edit Distance might be applied to intonation. Gussenhoven and Rietveld (1991), in a within-variety study, developed a fairly sparse measure of phonetic/phonological distance between intonation contours in the AM framework. Their treatment is much more comprehensive than Connolly's but the numerical scores assigned are quite simplistic. Their goal was not subtlety of measurement but to look for correlations between perceptual phonetic distance and meaning distance. It is possible that participants' perception of phonetic distance may have been wrongly influenced by meaning. Gussenhoven & Rietveld's work is a follow-on from Gussenhoven (1984). Gussenhoven (1984, ch.7), also in the context of comparison within a single variety, shows that phonetic distance in intonation is difficult to assess perceptually, even for trained phoneticians. Gussenhoven's findings reinforce the point that the problems Heggarty (2000a) identified as the Compatibility and Quantification problems are as relevant to intonation as they are to the segmental level. This is because the phoneticians had difficulty both in knowing which elements of the intonational contour should be compared with which and how to assess the degree of difference between different parts of the contours.

Nevertheless, in the broader experiment with naïve listeners which incorporated the above experiment with the phoneticians, Gussenhoven gave some evidence that listeners perceive the distance of contours in line with his three basic intonational contours and accompanying phonetic modifications. Gussenhoven (1984) and Gussenhoven and Rietveld (1991) argue that the Gussenhoven (1984) system is a more faithful representation of how people perceive phonetic and meaning distance in intonation than the Pierrehumbert (1980) system.[5] Ladd (2008, sec.4.2.1) also argues in favour of a Gussenhoven-style analysis as a way of capturing the intonational similarities, which he himself (Ladd, 1983) attempted to capture by the use of features (e.g. [delayed peak]).

---

[5]Herman and McGory (2002) have also assigned distance scores between intonational categories but in the context of understanding transcriber agreement and disagreement.

Hermes (1998) used ratings of auditory distance of pitch contours in order to develop visual aids in teaching intonation to the deaf or 2nd language learners. He also got ratings of visual distance of pitch contours in order to coordinate the two. He did not try to establish any correspondence between auditory distance and acoustic events though.

Patel et al. (2006) proposed a measure of distance incorporating duration, rhythm, and intonation. The Prosogram tool, which they use for the intonational measurements, would not be appropriate for analysing large amounts of data in the context of varieties, however. Patel et al. had the very different goal of making comparisons between language and music.

3.3 Nuclear statement rises in Belfast and Glasgow English

I apply the attempt to measure intonational distance to the very specific phenomenon of nuclear statement rises in a group of English varieties known as 'Urban North British' (UNB) (Cruttenden, 1995, 1997). This group includes Belfast and Glasgow English but not other Scottish English varieties (Brown et al., 1980; Aufterbeck, 2003). The nuclear rises spoken on statements in these varieties have attracted attention for some time but very little empirical work. These rises have been of interest because they contrast with the view that statement intonation generally features a nuclear *fall* in pitch, while *questions* contain rising pitch (Bolinger, 1978; Haan, 2002; Gussenhoven, 2004). So are these UNB rises really rises in the sense of typical question rises?

By 'nuclear rises' I restrict the examination of intonation to the nuclear tone, the part of the intonation contour beginning with the most prominent stressed syllable (nuclear syllable or 'nucleus'), which is the last pitch accent in the Intonational Phrase (IP) (Pierrehumbert, 1980) or the intermediate phrase (Pierrehumbert and Beckman, 1988). The nuclear tone continues over all subsequent unstressed syllables ('tail') until the end of the Intonational Phrase (IP). The nuclear tone is a fundamental unit of the British tradition of intonational analysis (Crystal, 1969; Cruttenden, 1997) and has been incorporated into AM analysis (Ladd, 2008, sec.4.1.1). This gives a well-defined counterpart to the use of cognates in segmental phonetic distance measures. Comparing nuclear tones addresses the Compatibility Problem (Heggarty, 2000a) of ensuring that we compare like with like.

Jarman and Cruttenden (1976) represents one of the earliest and most influential papers on the intonation of Belfast English, though there is some slightly earlier work by Knowles (1975) on Liverpool English (see also Knowles, 1978, 1984).[6] Cruttenden (1997) describes four different variants of UNB rises using the analysis of the British tradition: 'rise','rise-plateau', 'rise-plateau-slump' and 'rise-fall'. The 'rise-plateau' features a rise followed by a levelling off of pitch. The 'rise-plateau-slump' features a further falling off of pitch. Cruttenden (1997, p.133ff.) claims that the 'rise' is dominant in Glasgow English, while the 'rise-plateau-slump' is common in Belfast, Liverpool and Tyneside English. Ladd (2008) follows this with UNB rises now cast in the AM framework. Mayo et al. (1997) deal with Glasgow English in the AM framework but only include short observations on the phonetic realisation of the Glasgow rise. Vizcaino-Ortega (2002) examined the phonetic realisation and phonological representation of just four y/n questions from the HCRC Map task corpus of Glasgow English (Anderson et al., 1991). Wells and Peppé (1996) examined turn-taking in spontaneous Belfast English speech from the 1970s. The only systematic work involving acoustic analysis is that of the Intonational Variation in English (IViE) project (Grabe et al., 2001; Slater, 2008) and subsequent OXIGEN project (Grabe et al., 2008), combined with Lowry's work specifically on Belfast English (Lowry, 1997, 2002*a,b*). The IViE project included Belfast English (along with a few other UNB varieties) but not Glasgow English.

Comparisons of UNB rises with the intonation of other varieties of English in IViE project primarily involved:

- contrasting the phonological labelling of statement and question intonation in an AM approach using IViE's own labelling system (Grabe, 2002; Grabe and Post, 2002);
- contrasting the percentage frequencies of these rises in different varieties (Grabe, 2002; Grabe et al., 2005).

Neither of the two papers dealing with phonetic analysis (Nolan and Farrar, 1999; Grabe et al., 2000) had the goal of comparing the phonetic realisation of the UNB rises in the different UNB varieties, though these papers did examine variation among the varieties in other ways.

---

[6]The earliest reports of the UNB phenomenon appear to come from Tyneside English. See Bolinger (1978) and references therein.

Separately, work on the Irish language has also shown statement nuclear rises in the northern Donegal dialect, in contrast to the nuclear statement falls found in the Southern dialects (Dalton and Ní Chasaide, 2005; Dalton, 2007). The assumption so far is that these Irish rises are part of the same phenomenon as the UNB English rises.

3.3.1 Possible origins

The unanswered question is where these rises have come from. Dalton & Ní Chasaide explored the possibility that these rises were actually phonetically modified versions of falls. This possibility forms a major component of the theoretical proposals I develop in section 3.4 below. They realised that quite complex phonetic changes would have to have happened for a Connaught nuclear fall to become a Donegal nuclear rise, however.

Prior to Dalton & Ní Chasaide's work, Cruttenden and Knowles had put forward the view that the UNB rises emerged due to the influence of Irish on English and/or movements of Irish English speakers to parts of the UK. Dalton & Ní Chasaide continue this view, explicitly stating that they assume the UNB rises in English show the effects of transfer from the Donegal Irish rises. It is not within the scope of this thesis to explore the possibility of transfer or substratum influence from Irish much further. However, there is little evidence to favour the direction being from Irish into English as opposed to the other way around. Cruttenden (1995) argues more for an "urban spread" (p.160) than for the influence of Irish.[7]

Bolinger (1978, p.510) considered that UNB rises had derived from question intonation and that they represented the same phenomenon as High Rising Terminals (HRTs). Bolinger proposed that UNB English speakers started using question rises on statements when these statements had an element of questioning meaning e.g. to express uncertainty in what they were saying. However, the questioning meaning became redundant over time and speakers started using rises on statements much more widely. HRTs also involve rising intonation on statements as well as questions (Cruttenden, 1997, p.129ff.). First associated with Australian and New Zealand English (McGregor, 2005; Warren, 2005) and separately, U.S. English (McLemore, 1991), there are now reports that HRTs have

---

[7]Cruttenden (1995) suggests that UNB rises may have been present in Tyneside English before any influx of Irish English speakers. He bases this on a small anecdote from a Tyneside speaker in the 18th Century, quoted at the beginning of this chapter.

spread to the southern UK (Cruttenden, 1997; Shobbrook and House, 2003).[8] Fletcher et al. (2005) include some treatment of alignment from the IViE UNB data and Antipodean rises but do not formally compare them.

This section demonstrates that there are key gaps in our knowledge of the UNB rises. First of all, there is simply an absence of instrumental phonetic analysis of these rises. Unlike Dalton & Ní Chasaide's brief phonetic comparisons of Connaught falls and Donegal rises in the Irish language, there have been no such comparisons between UNB rises and falls in other English varieties. It is also unclear how similar UNB rises actually are to question rises in Standard English varieties. Further, it is unclear whether there are systematic phonetic differences between statement and question intonation in UNB varieties (though see the brief treatments in Ladd (2008, p.127) on Glasgow English and in Grabe (2002) on Belfast English), since both statements *and* questions use nuclear rises. Despite the inclusion of a few UNB varieties in the IViE project, Glasgow English was not included. Therefore, systematic comparisons of Belfast and Glasgow English intonation are lacking. Before tackling these gaps empirically, I develop theoretical arguments on intonational change.

3.4   Theoretical proposals on intonational change

First I would like to distinguish between the spread of an intonational change and the actual process of phonetic change in intonation. A familiar example of the spread of a change is the increasing use of High Rising Terminals (HRTs) in the English speaking world (section 3.3.1 above).

By contrast, phonetic change in intonation may involve subtle gradual phonetic changes to a given contour type in an individual language variety. I wish to focus primarily on changes of the alignment of local $f0$ maxima and minima with respect to the segmental content of the nuclear tone. The possibility of alignment change in intonation is directly influenced by the study of alignment differences between the lexical pitch accent I and II in Swedish and other Scandinavian languages. Dalton and Ní Chasaide (2005) and Dalton (2007) explored the "Re-alignment hypothesis" in the Irish language. This hypothesis stated that the nuclear and prenuclear statement falls in the Cois Fharraige Connaught dialect of Irish represented the same underlying contour as the nuclear and prenuclear rises in the Gaoth Dobhair Donegal dialect. The difference in surface realisation

---

[8]This rising intonation pattern has also been termed 'Uptalk'.

was a much later aligned peak in the Donegal dialect. The peak in the Connaught dialect was aligned with the stressed syllable with a fall across subsequent syllables. In the Donegal dialect, the peak was aligned with an unstressed syllable following the stressed syllable. This late alignment of the peak was accompanied by a rise beginning around the end of the stressed syllable. Dalton & Ní Chasaide rejected the hypothesis when they discovered that the Donegal contour was not simply a replica of the Connaught contour, just with later alignment of the peak.

Dalton & Ní Chasaide did not reject the more general hypothesis that there was historical change which led to falls being realised as rises in the Donegal dialect. They concluded that the amount of change that would have had to happen would be too great to consider the Cois Fharraige alignment of H and Donegal alignment of H as 'allophones' of an underlying intonational 'phoneme'. This is a perfectly valid assessment following from their analysis, apart from the thorny issue of how similar the alignments would have to be regarded as variants of each other.[9]

I take up where Dalton & Ní Chasaide left off and develop theoretical proposals for how "Re-alignment" of the peak could lead to other phonetic changes. These other phonetic changes would make the "Re-aligned" intonational contour phonetically distinct from its predecessor. This might explain why the Donegal contour was not simply a replica of the Connaught contour.

3.4.1 Late alignment of peaks

The peak (H) in nuclear falls is often aligned at the very end of the stressed vowel or even beyond it (Silverman and Pierrehumbert, 1990; Arvaniti et al., 1998; Gussenhoven, 2007).[10] The trend for late alignment of H is also found in languages with lexical tone (Akinlabi and Liberman, 2001; Silverman, 1997; Xu, 1998). There may be perceptual and articulatory reasons for this (Silverman, 1997; Xu, 2005; Hyman, 2007). One early study showed that listeners perceived the pitch on an entire vowel to be the pitch produced in the last third of that vowel (Rossi, 1971). Another finding is that for a specified pitch interval, speakers could produce a fall quicker than they could produce a rise (see the studies cited in Silverman (1997) and the more recent work by Xu and Sun (2002)). Xu (2001) accounts for the late alignment of H targets by arguing for a delay between

---

[9]This is a conundrum very much akin to Laver (1994) on [h] and [ŋ] in English.

[10]Dalton & Ní Chasaide examined prenuclear and nuclear falls and rises in the Irish language. I restrict my own proposals and later experimental analysis in chapter 4 largely to nuclear tones.

the "neural command" to produce the end of a rise and the action of the vocal cords in articulating it.

3.4.2 Rightward H drift

An extension of the trend for late alignment of H is the cross-linguistic tendency for H to drift[11] rightwards, particularly with an increase in the number of following unstressed syllables (Steele, 1986; Silverman and Pierrehumbert, 1990; Silverman, 1997; Dalton, 2007; Hyman, 2007). H often drifts beyond the accented syllable itself. Gussenhoven (2007, p.420) also invoked this synchronic tendency as a potential motivator of diachronic change in intonation (citing Dalton & Ní Chasaide's work, and van Leyden and van Heuven (2006) on Orkney and Shetland English). Dalton & Ní Chasaide link late alignment of H to H drift, suggesting that late alignment of H predisposes it to drift further. In Cois Fharraige Connaught Irish, prenuclear H was timed around the right edge of the syllable. This was one of the few pieces of support Dalton & Ní Chasaide found for a "Realignment" of the H accounting for the difference between Connaught falls and Donegal rises.

H drift over time has been a prominent account of alignment change in Scandinavian languages. Hognestad (2006), using archive data from Stavanger Norwegian in 1927 and 1970, noted that the earlier data showed earlier alignment of H for accent I. He made use of the [delayed peak] feature from Yip (2002, p.8ff) and ultimately Ladd (1983), suggesting a move from the first to the second mora from the 1927 to the 1970 data. Hognestad's own contemporary data shows an even later alignment of the H than in 1970 datasets. Of course, these 3 datasets are very different to each other and would have had very different test materials for the speakers involved. Alongside this historical account, Hognestad gives examples of contemporary varieties of Norwegian that fit each of these three alignment patterns. Therefore, he believes that contemporary dialectal variation can be ultimately mapped onto stages of historical change. For accent II, Hognestad also argues for rightward movement of H, but it is even less clear here whether there is directly comparable historical data to support this.

Both Bye (2004) and Hognestad (2006) also invoke rightward movement of H to account for the development of the lexical pitch accent distinction between accent I and II in Scandinavian. Drawing on earlier similar explanations of the

---

[11]I do not use this term in the sense of evolutionary drift. Drift in the context of my account simply refers to small gradual changes in the alignment of H.

Hypothetical path of change

Rightward peak drift

σ σ **(nuclear syllable)** σ
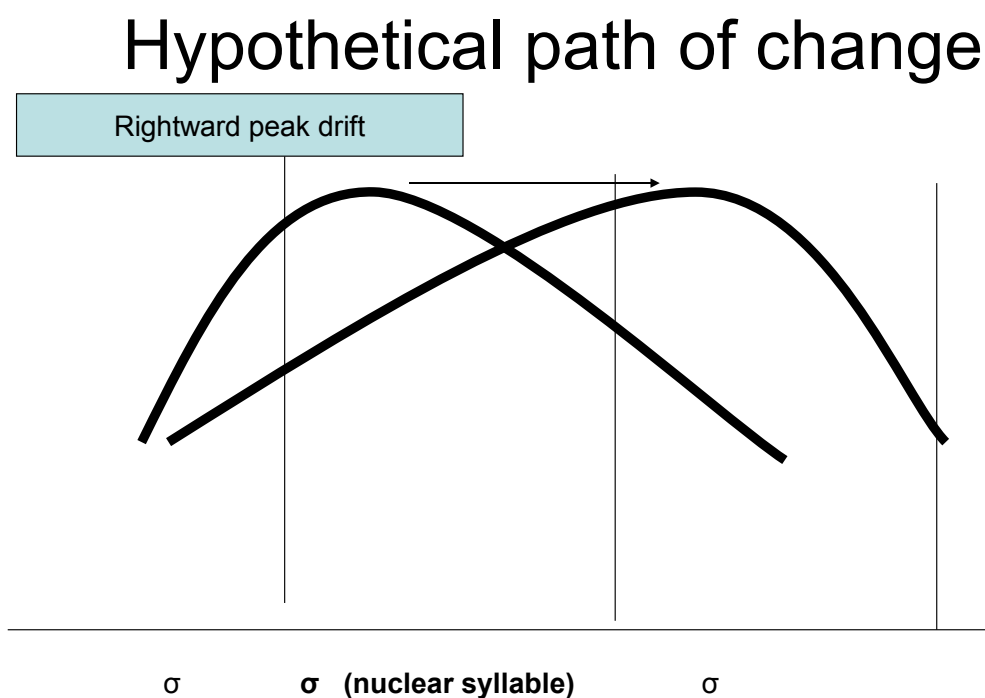
Figure 3.2: Diagram illustrating rightward peak (H) drift in the hypothetical path of alignment change. The Greek letter 'sigma' indicates 'syllable'.

Swedish lexical pitch accent contrast, Cruttenden (1997) (with reference to Knowles (1975) on Liverpool English) suggests that a [delayed peak] feature and/or peak spreading could account for the difference between RP statement falls and the UNB 'rise-plateau-slump'. Rightward spreading of H lexical tones is also well-known (Silverman, 1997; Hyman, 2007). The parallels between intonation, lexical pitch accent and lexical tone allow me to draw on existing work on the latter two to develop ideas about change in intonation.

Rightward H drift is problematic for the "Re-alignment" hypothesis in Irish dialects. Neither the Cois Fharraige dialect nor the Donegal dialect actually displayed the common tendency of H to drift rightwards with an increase in the number of following unstressed syllables (Ní Chasaide and Dalton, 2006; Dalton, 2007; Dalton and Ní Chasaide, 2007). If H does not drift rightwards synchronically in the Cois Fharraige or Donegal dialects, then why would it do so diachronically? Of course, it is unclear what either of these dialects may have done in the past. H may have been more variable at an earlier stage which led to change, or it may not have been. More recent work on Donegal English does

in fact show H drifting rightwards with an increase in the number of following unstressed syllables (Kalaldeh et al., 2009).

An extremely tentative alternative motivation for alignment change is speech style. Hognestad (2006) suggests that more rightward alignment of H may be linked with more colloquial styles but this is based on informal observations of the speech of just two speakers.

3.4.3 Rightward L drift

I have focused on rightward H drift, but what about rightward L drift? Previous studies have not examined the alignment of L valleys as rigorously as H peaks. Some studies though have shown the alignment of H to be more variable under different prosodic conditions in a synchronic context than L (Caspers and van Heuven, 1993; Prieto et al., 1995; Arvaniti et al., 1998; Ladd et al., 2000; Ladd and Schepman, 2003; Prieto and Torreira, 2007). Prieto et al report that L valleys are not as affected by the preceding (left-hand) prosodic context as H peaks are by the following (right-hand) prosodic context. However, very few researchers have examined if L is affected by the right-hand context in the same way as H. Rightward movement of the whole rise is what van Leyden and van Heuven (2006) argue to explain the difference between Shetland English and Orkney English. However, in the details of their analysis they actually focus on the rightward movement of the H and not of the preceding L. Xu (1998, p.201-2), working on Mandarin, also claims the whole rise moves to the right as syllable duration increases. Hyman (2007) (with reference to his earlier work) found that cross-linguistically, H lexical tones were more likely than L tones to spread rightwards. Perhaps this could be a simple consequence of the greater numerical frequency of H tones than L tones in lexical tone languages. Nevertheless, Yip (2002) (cited in Bye (2004)) does allow valley delay in lexical tone languages and Bye (2004) draws on this to account for the difference between two dialect types in Swedish.

I have focused on the trend for rightward alignment change in intonation, lexical pitch accent and lexical tone. From previous studies, the dominant direction of alignment change does indeed seem to be rightwards. To present a fuller picture, I should add that leftward alignment changes have also been reported (Engstrand and Nyström, 2002; Hualde, 2007; de Boer, 2009). These include the historical account of dialectal differences in the 1 peaked accents in Scandinavian.

3.4.4 *Gradual change becomes categorical*

If a H drifts too far to the right, listeners may no longer perceive it as associated to the nuclear syllable. This may lead to perceptual reorganisation and phonological change (see Asu and Nolan, 2007, p. 576-7, 582-3), such that the H is now perceived as belonging to a following unstressed syllable, perhaps akin to metathesis in segments. Misperception as a motivating factor in phonological change (including types of metathesis) is central to Blevins (2004, 2007), though it is not the full story there.

Dilley (2005, p.61) also proposed that alignment of $f0$ maxima and minima are tightly linked in perception. Her perception experiments revealed categorical perception once the alignment of H moved across a syllable boundary, and her production imitation task seemed to complement this finding (p.102). Pierre-humbert and Steele (1989) also argued that in *production* US speakers made a categorical distinction between early and late aligned H targets. This was not clear cut for all speakers, however. Following House (1990), Dilley found that the best point for discriminating H targets was often at vowel onsets or other places where sonority was increasing very fast. This is different to the common finding that H tends to align towards the end of the vowel. It is possible that several articulatory and perceptual factors interact in the coupling of the pitch/$f0$ events to the segmental material.

3.4.5 *Phonetic changes to L*

Arvaniti et al. (2000), with reference to Grice (1995, p.189-90), also talk of Hs being more easily perceived than Ls unless L is clearly different in scaling from the surrounding tonal targets. Indeed there may be a trend for L targets to be level stretches or broad troughs (Arvaniti et al., 2006; Grice et al., 2000; Lickley et al., 2005), in contrast to H targets which are often realised as sharp peaks. However, the realisation of a tone as a flat stretch instead of a single $f0$ point does not fit easily with AM theory (Arvaniti et al., 2006). Knight and Nolan (2006) show that H targets may often be flat stretches ('plateaux') too.

One of the major phonetic differences between the Connaught fall and the Donegal rise in Dalton & Ní Chasaide's work was that the pitch range between the L and H before the main fall in the Connaught contour was narrower than the

Hypothetical path of change

σ   σ (nuclear syllable)   σ
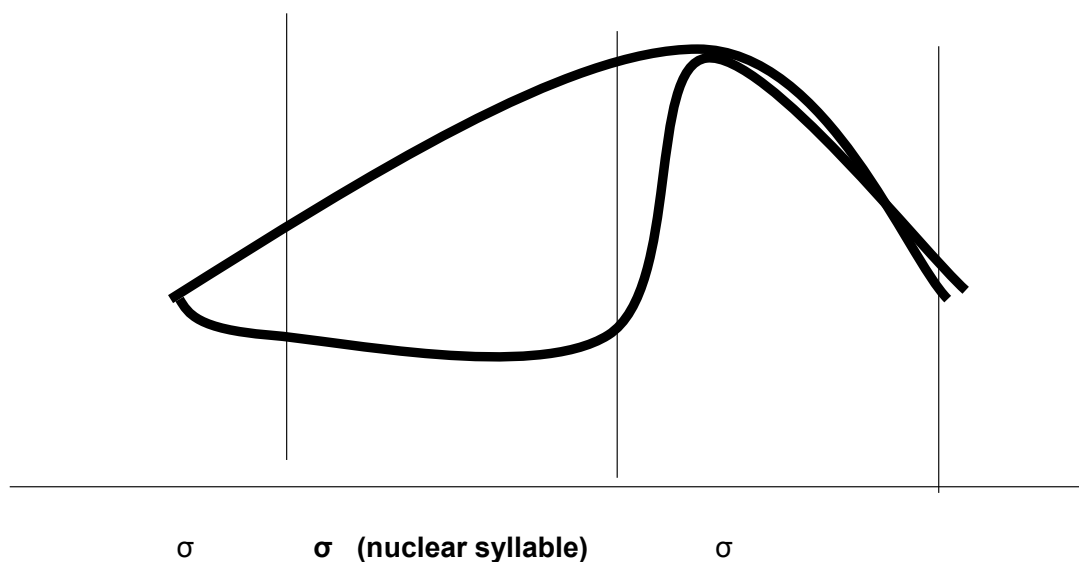
Figure 3.3: Diagram illustrating the potential phonetic changes to L that could result from H drifting beyond the nuclear syllable. The lower thick black line represents a nuclear intonational contour in which the L tone associated to the nuclear syllable has become a broad dip.

pitch range of the L and H delimiting the Donegal rise.[12] If rightward movement of the H in the Connaught contour led to the formation of the Donegal contour, then we might expect the L in the Donegal contour to be scaled lower than the L in the Connaught contour. The Donegal L also fitted the pattern of being a rather flat stretch (Dalton, 2007). In the delayed peak accents of Saaremaa Estonian, the accented syllable also has lower pitch than on the non-delayed peak accents of Standard Estonian (Asu, 2006).

3.4.6   The effect of L on H

The need for L targets to be broad dips rather than sharp troughs/turning points may have further implications for the alignment of the following H. To avoid tonal crowding of the L and H, the H may need to drift even further to the right. There are precedents for this proposal. Arvaniti et al. (2006) found that when the

---

[12]Dalton and Ní Chasaide (2005, p.452) has a diagram illustrating the difference.

Hypothetical path of change

σ   σ **(nuclear syllable)**   σ
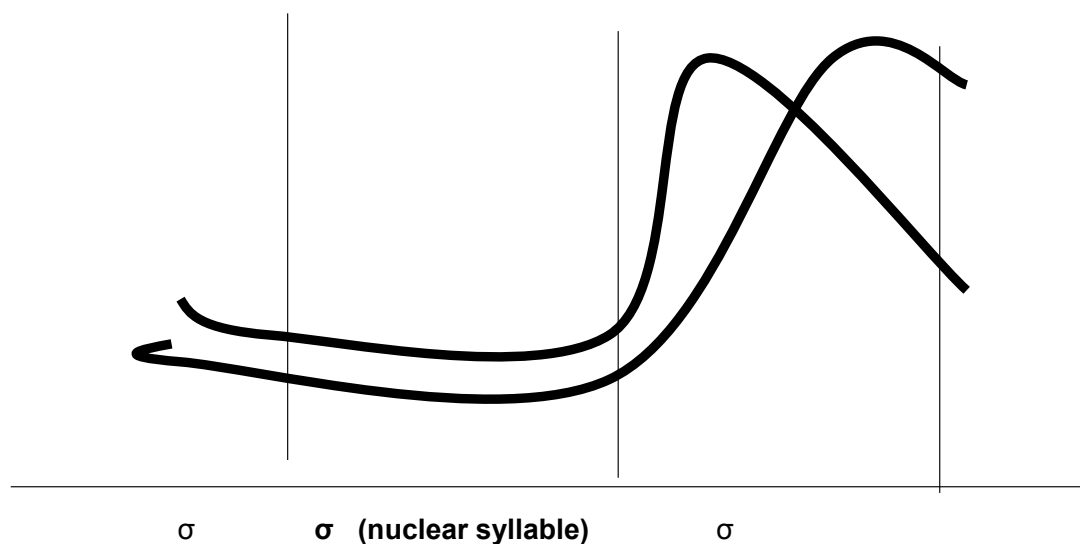
Figure 3.4: Diagram illustrating the potential phonetic effects of L on H. The lower thick black line represents a nuclear intonational contour in which the H in the final syllable has moved even further to the right following the presence of L as a broad dip on the nuclear syllable itself.

nuclear syllable is the final syllable in Greek, H occurs statistically further rightward than in other conditions. The authors argue that the vowel in this syllable needs to accommodate both the H and a preceding L target. The presence of the L then leads to a later alignment of the H (see also Riad, 1996; Kristoffersen, 2000, 2007; Peters, 2007).

Hognestad (2006) has a similar explanation of diachronic change in Stavanger Norwegian. In long compounds, the H of accent I is even further to the right of the post-stress syllable. The extent of the rightward movement, Hognestad argues, is due to the presence of an L target on the stressed syllable itself now. This requires the H to align later and leads to the formation of a plateau. It is not clear how tonal crowding could account for the presence of a plateau. Nevertheless, Hognestad's account is relevant both to Dalton & Ní Chasaide's work on Donegal Irish and to previous work on UNB rises in English. Dalton & Ní Chasaide argue that the Donegal rise finishes with a high plateau and that this

plateau is not present in the Connaught data. (From diagrams presented in their work, the distinction between peak and plateau does not actually appear to be as clear as this.) Belfast English rises have also been described as 'rise-plateaux' (Cruttenden, 1997; Grabe, 2002).

Avoidance of tonal crowding may also explain the finding of Nolan and Farrar (1999) that the presence of at least one unstressed syllable before the accented syllable made peak delay[13] much less likely than when there were no preceding unstressed syllables. Their data were from four English varieties in the IViE corpus (including Belfast English). The initial L target may need space to be realised so if there is no preceding unstressed syllable, the presence of the L on the accented syllable may result in a later alignment of the H. Nolan & Farrar examined initial pitch accents in IPs and not nuclear rises/falls, the latter being my interest. Nevertheless it is relevant to mention that Belfast displayed more peak delay than the three other varieties but less extreme peak delay (H on the 2nd following unstressed syllable or later) than the Leeds and Newcastle data.

Mücke et al. (to appear) describe both the L and H of the German rise as being "in phase" with the accented vowel in their articulatory study. They claim that the H gesture has moved rightwards to make space for the L gesture beforehand and draw analogies between this and the rightward movement of a consonant when another consonant is inserted before it.

3.4.7 Truncation

The string of changes set off by the rightward movement of H may conclude with Truncation. The alignment of H may now be so far to the right that there is no longer enough 'room' to produce the final fall of the original contour (Gussenhoven, 2007). This might explain why in Donegal Irish and Belfast English, the contours are rises without much if any trace of a following fall. However, there is a danger of circularity. Gussenhoven (2004) first found evidence of truncation in the Tongeren variety of Flemish and uses that finding to argue that this variety must have a delayed peak. I have argued these points in reverse.

Grabe et al. (2000) found truncation in Belfast English data in the IViE project. Leeds English also displayed truncation but Grabe et al. do not discuss the clear phonetic differences between Belfast and Leeds. These differences indicate that truncation may not be an entirely unified phenomenon.

---

[13]Nolan & Farrar use the term 'lag' instead of 'delay'.

3.5   Approaching Chapter 4

The series of changes that I outline in this section 3.4 represent a plausible account of how a (nuclear) fall could change over time to become realised as a (nuclear) rise. In brief, these changes are:

1. "Re-alignment" of the H through rightward drift beyond the accented syllable;
2. "Re-alignment" of the preceding L onto the accented syllable;
3. Lowering of L scaling and broadening of the contour around L;
4. Further "Re-alignment" of H to avoid tonal crowding with L;
5. Truncation of final part of the original (nuclear) fall.

The next question is whether measurements of intonational distance could pick up on the connection between the "Re-aligned" and original contour despite these phonetic changes. This is where I turn to an empirical examination of nuclear rises and falls, first in Belfast and Cambridge English (Experiment I) and then in Belfast and Glasgow English (Experiment II). The results will demonstrate if there is a meeting point between the theoretical account and real data. The account of "Re-alignment" illustrated in this chapter becomes recast as the Alignment hypothesis in chapter 4.

CHAPTER 4

Intonational Experiments I and II

4.1 Introduction: Phonetic analysis

In this chapter I examine whether intonational distance measurements using contemporary data reflect the theoretical account of intonational change presented in chapter 3. My underlying interest is the phenomenon of nuclear statement rises in Belfast and Glasgow English. The key questions are where these rises have come from and whether intonational distance measurements can help to understand their origin.

The layout of this chapter is as follows: I first introduce the methodology and analysis of Experiment I, comparing Belfast English and Cambridge English, using data from the IViE (Intonational Variation in English) corpus (Grabe et al., 2001) (section 4.2 below). The results show a mismatch between the theoretical historical account (Alignment hypothesis) and the phonetic details of the Belfast English nuclear statement rises. They gave more support to an alternative hypothesis on the historical origin of the Belfast statement rise, the Transfer hypothesis (see section 4.5.4). I describe the need for a new larger-scale experiment to examine historical change and intonational distance more thoroughly (section 4.6). This new experiment is Experiment II, with specially designed materials and a few speakers from Belfast and Glasgow English.[1] This is the first step at addressing the dearth of formal comparisons of Belfast and Glasgow English intonation. I provide details of the methodological changes, analysis, results and discussion of Experiment II in section 4.7 below.

---

[1] I also included speakers from Edinburgh English but refer to analysis of the Edinburgh data only briefly.

My approach to measurements of intonational distance takes the nuclear tone as its unit and incorporates the two AM parameters of alignment and scaling as the primary parameters.

4.2 Experiment I: Analysis of IViE data

4.2.1 *Initial hypotheses*

I begin my approach to intonational measurement in the narrow context of the nuclear rises found on statements in Belfast English. I contrast these with statement nuclear falls in Cambridge English and with question rises in both Belfast and Cambridge English. Following Dalton & Ní Chasaide's lead (Dalton and Ní Chasaide, 2005; Dalton, 2007), I ask if these rises could have originally been more typical nuclear statement falls.

The Alignment hypothesis posits that statement intonation in Belfast and Glasgow English originally contained a standard nuclear fall but changed into a rise through rightward "Re-alignment" of the nuclear peak (Dalton and Ní Chasaide, 2005; Dalton, 2007). In such case, I expect more alignment distance between statement rises and question rises than between statement rises and statement falls.

The competing hypothesis is that UNB statement rises have derived from question rises and would thus be more similar to question rises than to statement falls. This is the Transfer hypothesis. Bolinger (1978) is the main scholar who has suggested this kind of transfer as an explanation for the UNB statement rise. Bolinger proposed that the transfer happened initially to statements that contained an element of questioning meaning e.g. 'are you listening'? Over time the statement rise lost the questioning meaning and was simply used as the general statement nuclear intonation contour (see also Cruttenden, 1997, p. 110).

A small degree of distance does not entail a shared origin, but the assumption of a connection between a small degree of distance and a common origin is a useful place to start.[2]

This thesis assumes that nuclear rises on statements in UNB varieties must represent some kind of change because it is not in line with the traditional description

---

[2]Some cognate forms in segmental phonology may in fact be phonetically distant (Campbell, 2003). There are a few well-known examples of this, such as Latin 'duo' and Armenian 'erku'.

of statements a displaying nuclear falls. However, it is possible that these varieties have always displayed statement rises and that there has not in fact been a process of Alignment or Transfer change.

4.2.2   Illustration of potential change

The way in which a statement nuclear fall like those found in Cambridge English could become a statement nuclear rise in Belfast English might be as follows:

1. The peak (H) is aligned within the nuclear stressed syllable. Importantly, there is a small valley (L) before H which rises from the previous unstressed syllable to the nuclear syllable. After H, the pitch falls noticeably (figure 4.1).

2. H drifts to the right towards the offset of the nuclear vowel. A more prominent and perceptible rise before H emerges, and there is still a final fall afterwards (figure 4.2).

3. H drifts even further to the right, well beyond the nuclear syllable. The nuclear syllable contains an L target, creating a rise from around the end of the nuclear syllable to the following unstressed vowel. There is no final fall after the H, perhaps because there is no room for it (see chapter 3) (figure 4.3).

These three examples show three different alignments of H. They may reflect a possible rightward movement of H over time, which has led to the occurrence of rises on statements in Belfast English.

4.2.3   Data

To examine the Alignment and Transfer hypotheses empirically, I used data from the IViE corpus (Grabe et al., 2001). Despite the lack of phonetic analysis of UNB rises by the IViE researchers themselves, the IViE corpus is still a major resource. Of the UNB varieties, I chose Belfast English. IViE project publications report a much higher percentage of statement rises in Belfast English than in Newcastle English (Grabe, 2002), so it made sense to use this richer database. Belfast is also closer geographically to Donegal. As I showed in chapter 3, the Donegal dialect of Irish also features nuclear statement rises and Dalton (2007) and Dalton and Ní Chasaide (2005) briefly explored the possibility that these rises might have derived from nuclear falls.

Figure 4.1: Spectrogram, waveform and *f*0 trace of the statement 'We arrived in a limo' spoken by a female Cambridge English speaker (IViE reference c-dec3-f1). This displays point 1 above. The IPA transcription in this and subsequent figures are broad.

I compared Belfast English data with data from Cambridge English (representative of Southern Standard British English (SSBE)).[3] I expected statements to feature nuclear falls in this variety in contrast to the nuclear statement rises of Belfast English. Questions in SSBE can have various intonation patterns, but a plain nuclear rise is one such pattern (Cruttenden, 1997). I also compare Cambridge question rises and Belfast question rises with Belfast statement rises.[4]

The IViE project includes data from five different speaking styles. I selected data from the Read Sentences and Read Passage styles. The list of Read sentences

---

[3]For information on how the IViE data were collected, see Lowry (2002*b*).

[4]Halliday (1967) thought that intonation was the key to distinguishing different sentence types in English. However, others have argued that a range of intonation patterns in questions is possible (e.g. Pike, 1945; Bolinger, 1989; Cruttenden, 1995).

Figure 4.2: Spectrogram, waveform and *f*0 trace of the same statement spoken by a female Belfast English speaker (IViE reference b-dec3-f3). This displays point 2 above.

(henceforth 'Sentences list') contained statements, wh questions, y/n inversion questions, declarative questions and coordination questions. These data had the potential for comparisons between statement and question intonation within each variety and between the varieties. The Read passage was a specially modified version of the fairytale Cinderella (henceforth 'Cinderella passage') and did not contain many questions. These data were therefore only appropriate for the comparison of *statement* intonation between the two varieties. Appendix D contains the IViE Sentences list and a transcript of the Cinderella passage.

All the Cambridge question data that I used took the form of nuclear rises with H aligned in the final syllable of the IP. This is not the only form of question intonation in the IViE Cambridge data. It was most frequent in declarative questions. A much smaller number of y/n questions and wh questions displayed this pattern
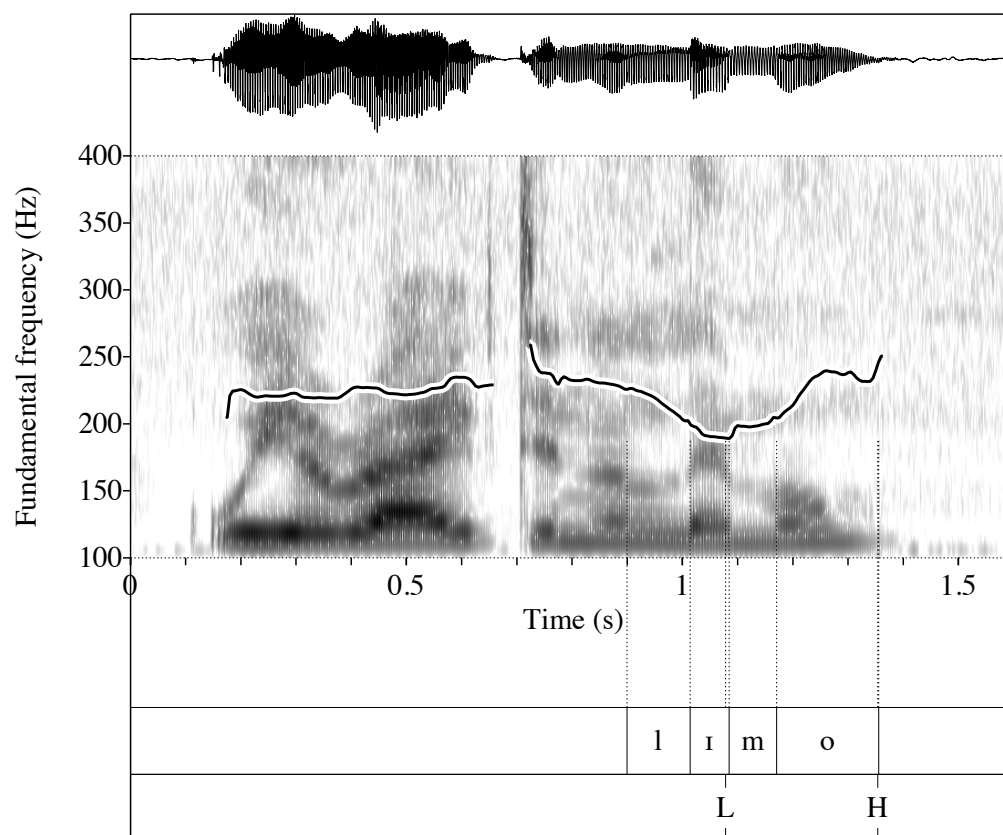
Figure 4.3: Spectrogram, waveform and $f0$ trace of the same statement spoken by a different female Belfast English speaker (IViE reference b-dec3-f1). This displays point 3 above.

despite Grabe et al. (2000, p.164) reporting it as a common pattern for standard British English y/n questions.

4.3 Methodology

4.3.1 Segmentation details

I used displays of the $f0$ contour, spectrogram and waveform with Praat software (Boersma and Weenink, 2009) for all acoustic analysis in this thesis. The first step was to delimit the IPs in the data. I relied on pauses and lengthening as cues to IP boundaries. de Pijper and Sanderman (1994) showed that there was good agreement among their raters for the placement of IP boundaries. In

the Sentences list, isolating the IPs was usually very straightforward. The sentences were short and thus usually constituted a single IP each. In some of the coordination questions, there was an IP boundary before the conjunction 'or' e.g. 'Are you growing limes, IP%[5] or lemons IP%?' In the Cinderella passage, IP boundaries often coincided with the punctuation marks of comma and full stop on the transcript, but not always. The next step was to demarcate the nuclear tone within each IP. The nuclear syllable is the most prominent stressed syllable in the IP and also the last stressed syllable in that IP. I marked the boundaries of this syllable and the following unstressed syllable (postnuclear syllable). I also segmented the vowel within each syllable. All of the final words in the IViE Sentences list were disyllabic with initial stress e.g. limo. In the Cinderella passage, I only analysed nuclear tones in which there was also just one unstressed syllable after the nuclear stress. I also marked the boundaries of the syllable (or monosyllabic word) immediately preceding the nuclear syllable (prenuclear syllable). This was usually unstressed. I adapted Praat scripts written by Pauline Welby to semi-automate the measuring procedures.

The first script included prompts for me to hand-segment the following elements:

1. prenuclear syllable (befnuc)
2. prenuclear vowel (bfv)
3. nuclear syllable (nuc)
4. nuclear vowel (v1)
5. postnuclear syllable (postnuc)
6. postnuclear vowel (pnv).

I mainly used the segmentation criteria in Turk et al. (2006). When I could not apply these critiera, I used the midpoints of formant transitions in marking segment boundaries between sonorant segments. I provide more specific details on segmentation in appendix D. The words in nuclear position in the Sentences list from the IViE project contained almost exclusively sonorant segments. The nuclear tones I analysed from the Cinderella passage also featured sonorant segments heavily. Sonorants do not cause as many segmental perturbations of $f0$ as obstruents. Sonorants, therefore, lead to a much smoother $f0$ contour which is more suitable for intonational analysis. The downside to this is that it is often

---

[5]% is the symbol used to mark an IP boundary in IViE's own intonational labelling system and in the ToBI system (Silverman et al., 1992; Beckman and Ayers Elam, 1997; Veilleux et al., 2004).

difficult to find reliable segment boundaries between two sonorant segments. I acknowledge that many cases of segmentation were tricky and accept that this has implications for the alignment results. This is because I measured alignment in milliseconds (ms) from the beginning or end of specific segments e.g. the onset of the nuclear stressed vowel.

The same script extracted the Pitch object in Praat and I applied smoothing of the $f0$ contour with the Praat default value of 10 bandwidth. This eliminated some though not all of the segmental perturbations of $f0$ and pitch tracking errors. From the smoothed Pitch object, the script located the $f0$ maximum (H1) within the nuclear tone automatically. The script also located the $f0$ minimum (L1) preceding the $f0$ maximum within the nuclear tone or within the prenuclear syllable.

The script marked all segment/syllable boundaries and $f0$ minima/maxima in a TextGrid accompanying each sound file (.wav). I checked all $f0$ minima and maxima by visual inspection. If an $f0$ minimum or maximum was clearly marked at a segmental perturbation or $f0$ error, I either changed the label location by hand or discarded the value if an appropriate alternative location was not obvious.

4.3.2   L 'elbows'

In addition to the true $f0$ minimum, I also marked the low turning point (or 'elbow') of the rise before the $f0$ maximum. This is the point at which the rise clearly accelerates (Xu, 1998, p.197). The true $f0$ minimum might be a rather random value in a long low stretch of pitch. Xu (1998); D'Imperio (2000) and Welby and Loevenbruck (2006) all discuss the difference between the 'elbow' marking the beginning of a rise and the true $f0$ minimum. I located and labelled the 'elbow' (EL) using the set of three scripts by Welby and Mary Beckman (and slightly adapted by myself) for marking 'elbows'. The first script is a Praat script and requires a stretch of the $f0$ contour to be already marked off by hand. The 'elbow' should be within this stretch. The first 'elbow' script extracts this stretch of $f0$ contour. The second 'elbow' script is an R script (R Development Core Team, 2010). This fits two lines to the extracted stretch of $f0$ contour. The 'elbow' is considered to be at the point of intersection of these two lines. The third script is a Praat script which inserts the location of this point in the

relevant TextGrid file. For further details on these scripts, see the accompanying notes at http://www.icp.inpg.fr/˜welby/PAGES/software.html (February 2011), D'Imperio (2000, p.92) and Welby (2003, p.69).

The disadvantage of using the line-fitting script is that rise turning points are often not as clear on the $f0$ trace as the work by Xu, D'Imperio and Welby & Loevenbruck would indicate. Shosted et al. (2006) found that turning points marked by algorithms did not always correspond to assessments by human raters as to the location of 'elbows' and in addition were affected by segmental perturbations of $f0$ (although Welby and Beckman's procedure performed better than the others tested). Worryingly, Ladd (p.c.) indicates that 'elbow' finding algorithms are not actually particularly objective procedures. Ladd indicates that the result of the line-fitting procedure is affected by how narrow or wide a section of $f0$ contour the researcher has marked off using the first 'elbow' script. On a subset of Belfast English data from Experiment II, I explored this and found it indeed to be a problem.

For Experiment I, I was not yet fully aware of these negative implications of using the line-fitting procedure. Therefore, I decided to use it to mark the L point at the beginning of the rise. If the line-fitting procedure marked L in an obviously inappropriate location, and if the true $f0$ minimum was appropriate, I used that instead. There was less of a problem with long high stretches at the end of the rise in these data. Part of the reason for this is that I only analysed data with one unstressed syllable after the nuclear stressed syllable. Therefore, I primarily used the true $f0$ maximum to mark the H ending the rise. For cases in which there was a clear turning point ending the rise before the true $f0$ maximum, I used the line-fitting procedure to mark this point.

In Experiment II, I abandoned the line-fitting procedure and used the $f0$ minimum and maximum to represent L1 and H1. I did not attempt to mark 'elbows' by hand as I believed it would be too subjective for a single researcher to do this. The $f0$ minimum was not just a random value in the middle of a long low stretch in these data.

4.3.3 *Extraction of values*

A further PRAAT script extracted all the values marked in each TextGrid file into a tab-separated text file. This included the following information:
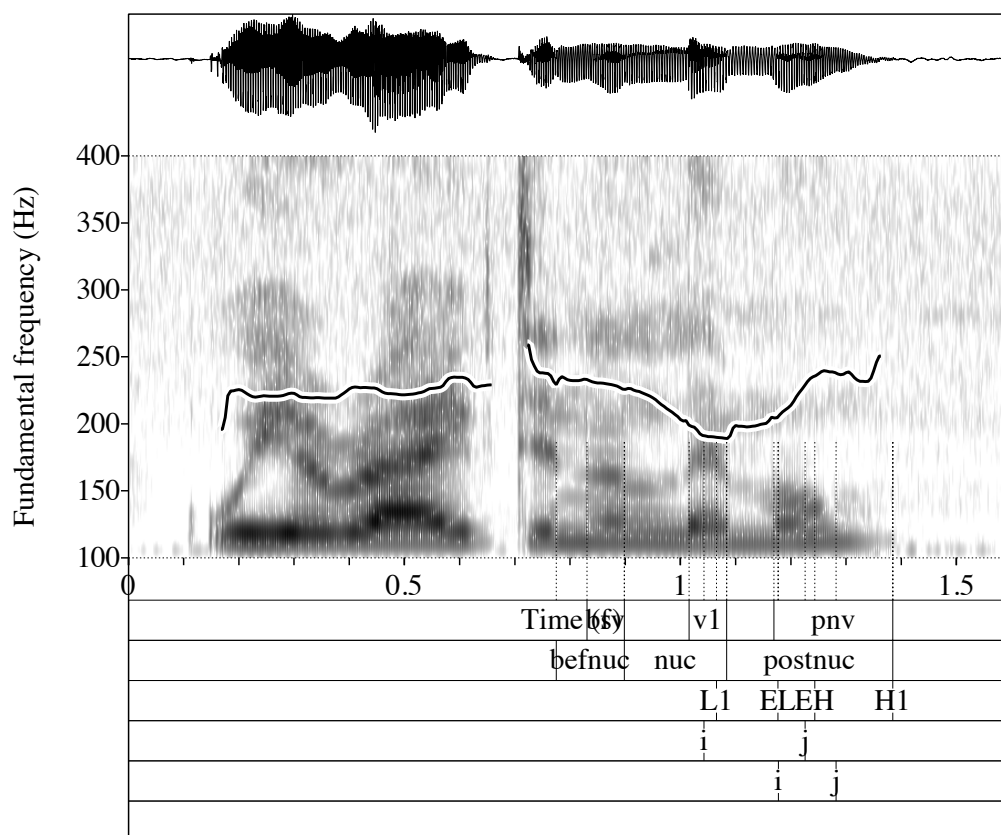
Figure 4.4: An example of a segmented sound file (IViE reference b-dec3-f1). The sentence is 'We arrived in a limo' and is spoken by a female speaker of Belfast English. The top part of the figure shows the waveform, spectrogram and $f0$ trace. The lower half of the figure shows the labelling from the Praat TextGrid file. The first tier (beginning from the top) shows segmentation of the relevant vowels (bfv, v1, pnv). The second tier shows segmentation of the relevant syllables (befnuc, nuc, postnuc). The third tier shows the location of the $f0$ minimum and maximum (L1, H1) and the 'elbows' (EL, EH). The fourth and fifth tiers show the delimited stretches of contour within which the R script located the 'elbows'. The last tier is a miscellaneous tier reserved for comments about each file e.g. notes about difficulty of segmentation.

1. the time point of each segment/syllable boundary;
2. the time point of the true $f0$ minimum (L1) and maximum (H1);
3. the time point of the low (EL) and if appropriate, high (EH) 'elbow';
4. the speaker's mean $f0$ and standard deviation in the utterances (in linear Hertz (Hz) and logHz);
5. the $f0$ values for L1, H1, EL and EH (in linear Hz and logHz);

6. the $f0$/pitch excursion between various combinations of L1, H1, EL and EH (in linear Hz);

7. alignment measurements in milliseconds (ms) for L1, H1, EL and EH from various segmental landmarks.

I did not make any measurements of speech rate in this initial analysis of the IViE data.

The extracted values were directly relevant to the parameters of interest.

4.3.4 Parameters

- Alignment of the beginning of the rise (L1 or EL)
- Alignment of the end of the rise (H1 or EH)
- Scaling of the beginning of the rise (L1 or EL)
- Scaling of the end of the rise (H1 or EH)
- Magnitude of the rise ($f0$ excursion between L1/EL and H1/EH)

For the Cambridge statement falls, the alignment of L1/EL refers to the valley preceding the peak in the nuclear stressed syllable, not the final low point at the end of the fall.

There are two main ways of measuring alignment:

1. absolute alignment in milliseconds from a given segmental landmark;
2. alignment as a *proportion* of the duration of a given vowel/syllable.

Knight and Nolan (2006) examined both types of alignment and considered that proportional alignment better accounted for the data than absolute alignment (see also Steele, 1986; Silverman and Pierrehumbert, 1990), though the choice is still an unresolved methodological issue (Prieto et al., 1995; Atterer and Ladd, 2004; Braun, 2006; Schepman et al., 2006). Admittedly, absolute alignment may be affected by the confound of different syllable durations. In Experiment I, I also explored both types of alignment. I found less variability across a set of utterances from a given sentence type in a given variety when I used absolute alignment. For example, in the set of Belfast English statements from the Cinderella passage, there was less variability across utterances when measuring the alignment of L as the absolute distance from the offset of the nuclear vowel than

when measuring it as a proportion of the duration of the nuclear vowel. Therefore, I report only absolute alignment measurements in the results. In Experiment II, I refer only to proportional alignment in the results.

In Experiment I, I refer only to the $f0$ excursion in relation to scaling measurements. In Experiment II, I prefer to use scaling values for L1 and H1 normalised by the speaker's mean $f0$. These values are more speaker-independent than the $f0$ excursion. I henceforth use the symbol L as a cover symbol for both L1 and EL and likewise use the symbol H for both H1 and EH.

4.3.5 Usable IPs

The total number of usable IPs analysed was 231 IPs in the Belfast data (139 statements, 92 questions) and 115 IPs in the Cambridge data (67 statements, 48 questions) giving a combined total of 346 IPs. Further details on the number of utterances (IPs) I analysed in both sentence types and in both English varieties are given in appendix D. In this initial study, the number of IPs in each category was unequal as was the number of male and female speakers, as was the number of utterances from each speaker.

4.4 Results

I used R for all statistical procedures reported in this thesis. In Experiment I, there was one independent variable: Category. This had five levels: Belfast statements (Sentences list), Belfast statements (Cinderella passage), Belfast questions, Cambridge statements, Cambridge questions. The dependent variables were the alignment and $f0$ excursion parameters listed in section 4.3.4 above.

4.4.1 Alignment

The alignment of L and H was very similar between the Belfast statements and all of the question rises. L was aligned after the offset of the nuclear vowel. H was aligned well beyond the onset of the postnuclear vowel. The mean alignment values for L and H in each variety and each sentence type are given in appendix D.

Using One-Way ANOVAs,[6] there were no significant differences in the dependent variables of alignment of L or of H between the Belfast statements, Belfast questions and Cambridge questions.

- L alignment, $F(3, 252)$[7] = 1.586, n.s.
- H alignment, $F(3, 275)$ = 0.5831, n.s.

By contrast, both L and H in the Cambridge statements were aligned much earlier. The L in the Cambridge statements was aligned shortly after the onset of the *prenuclear vowel* and H in the Cambridge statements was aligned shortly after the onset of the *nuclear* vowel.

4.4.2 f0 excursion

Subtracting the Hz value for L from the Hz value for H gives a measure of the $f0$ excursion between these points in each IP. This displays how great the rise actually is. The widest excursion was in the Cambridge questions, followed by the Belfast questions, Belfast statements (Sentences list), Cambridge statements and Belfast statements (Cinderella passage). The means are in appendix D. Using a One-Way ANOVA, there was a significant effect of Category on the $f0$ excursion ($F (284, 4)$=6.1894, $p < 0.001$). Post-hoc t tests with the Bonferroni correction are presented in appendix D. The most relevant findings from the post-hoc tests were:

1. Cambridge questions had the widest $f0$ excursion, significantly different from the Cambridge statements and the Belfast statements, but not the Belfast questions.[8]
2. Cambridge statements did not have a significantly different $f0$ excursion from Belfast statements or questions.
3. Belfast statements (Cinderella passage) and Belfast questions differed in $f0$ excursion, though Belfast statements (Sentences list) and Belfast questions did not.

---

[6]I use more advanced statistics, Linear Mixed Effects models, for the analysis of the data from the larger-scale Experiment II (section 4.9.1 below).

[7]The second term within the degrees of freedom parentheses refers to the number of usable tokens of alignment of L or H minus one. I had to discard various L and H alignment values if L and/or H were marked at obvious segmental perturbations of $f0$ or at $f0$ tracking errors. This explains why the degrees of freedom are not equal for L and H above.

[8]The data in Grabe et al. (2000) show Cambridge questions with larger $f0$ excursions than Belfast questions, though Belfast questions did not have unambiguously smaller excursions than Newcastle and Leeds questions.

4.4.3 Potential tonal crowding

In chapter 3, I raised the possibility that the alignment of L on or after the nuclear syllable might affect the location of the H. Specifically, it might result in an even later alignment of H. I tested this with Correlation tests for the Belfast statements. The results are in appendix D. The non-parametric Spearman correlation showed a significant (positive) correlation between the alignment of L and the alignment of H in the Sentences data only. This gives some indication that a later alignment of L may be accompanied by a later alignment of H, in line with my theoretical proposal in chapter 3. However, this is not conclusive.

4.5 Discussion

4.5.1 Alignment

This initial analysis showed a clear difference in the alignment of L and H between Belfast and Cambridge English statements. The Alignment hypothesis posits that H in Belfast has moved rightwards. In the alignment of both L and H however, Belfast statements are clearly more distant from Cambridge statements than from questions in either variety. H in the Cambridge statements was aligned within the nuclear vowel. Notably, it was not actually aligned towards the end of the nuclear vowel, different to the trend I discussed in chapter 3. H in the Belfast statements and in the Belfast/Cambridge questions was aligned towards the end of the postnuclear vowel. The preceding L in the Cambridge statements was aligned in the prenuclear vowel, whereas the L in the Belfast statements and in the Belfast/Cambridge questions was aligned around the end of the nuclear vowel.

Though the Belfast statements do indeed have a more rightwardly aligned L and H than the Cambridge statements, the degree of alignment distance between the Belfast statements and the Cambridge statements gives one indication that Belfast statement nuclear intonation may not be a derivation of nuclear statement intonation typical of Southern British English. The brief testing of tonal crowding (section 4.4.3) also did not clearly show that a more rightward L alignment was consistently accompanied by a more rightward H alignment in these Belfast data. In this initial study, all target nuclear syllables had just one following postnuclear unstressed syllable. What would happen if there were further unstressed syllables? This would help to evaluate potential links between the
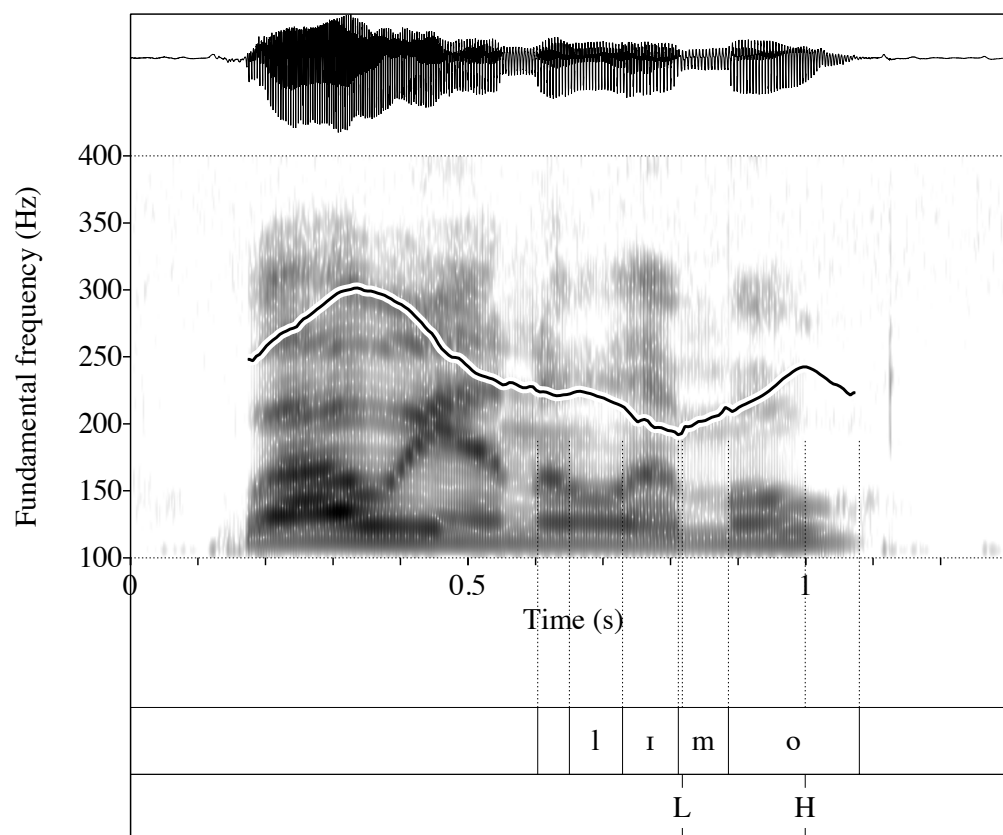
Figure 4.5: Spectrogram, waveform and *f*0 trace of the wh question 'Why are we in a limo?' spoken by Belfast English speaker f2 (b-whq3-f2). Compare with the statement 'We arrived in a limo' spoken by Belfast English speaker f1 in figure 4.3 above.

present alignment of the Belfast English H and potential rightward movement of this H over time.

The alignment of L and H in these Belfast English data appear very similar to nuclear statements in Donegal Irish (Dalton and Ní Chasaide, 2005; Dalton, 2007) and Donegal English (Kalaldeh et al., 2009). This furthers the assumption that the same phenomenon is taking place in both languages. This contrasts with Mayo et al. (1997), who reported that Glasgow rises begin early, with the L even earlier than the accented syllable itself. The alignment of L and H in these Belfast data may in fact be similar to that of HRT rises (Fletcher et al., 2005; Warren, 2005). Experiment II introduces direct comparisons between Belfast and Glasgow English

Figure 4.6: Spectrogram, waveform and *f*0 trace of the wh question 'Why are we in a limo?' spoken by Cambridge English speaker f4 (c-whq3-f4). Compare this with the Belfast question (figure 4.5 above), the Belfast statement (figure 4.3 above) and the Cambridge statement (figure 4.1 above).

nuclear intonation and also makes more indirect comparisons between these varieties and HRTs.

4.5.2 f0 excursion

It is not surprising that the Cambridge question rises should have had the widest pitch excursion. This is because of the recurring finding that questions tend to have a wider pitch excursion and/or a higher H than corresponding statements (see chapter 3). In this initial IViE analysis, the Belfast questions did not have an unambiguously larger pitch excursion between L and H than the Belfast statements. This may appear to go against the recurring finding mentioned above.

Though the Cambridge statements did not have a significantly different pitch excursion to the Belfast statements, this does not really give support to the Alignment hypothesis that Belfast statement intonation has derived from statement intonation like that found in Cambridge English. Existing work shows that there can be alignment differences between varieties but there is very little indication of the range of variability in $f0$ excursion that might be expected between varieties. Existing work rather focuses on within-variety variability in $f0$ excursion e.g. differences between statements and questions.

Even still, the overall connection between Belfast statements and Belfast questions is closer than the connection between Belfast statements and Cambridge statements. This gives more support to the Transfer hypothesis, that statement and question intonation in Belfast English may be part of a unified phenomenon. There was a clear difference in $f0$ excursion between Belfast statements and Cambridge questions, however. This might suggest that transfer of question intonation onto Belfast statements is not the only explanation of the Belfast statement phenomenon. The wider $f0$ excursion in the Cambridge questions might also be a consequence of the majority of the Cambridge questions being declarative questions (section 4.2.3 above), as these may display an even wider excursion or higher scaling of H than other question types (Grabe, 2002; Haan, 2002). Both Belfast statements and Belfast questions tend to jump down in pitch onto the stressed syllable, whereas neither of the Cambridge contours displayed this. For question intonation typical of Cambridge English to change into Belfast statement intonation, a reduction in $f0$ excursion would be necessary. The Belfast rises of normal hearing participants in Rahilly (1991, p.107-8) showed an average excursion of 54hz, much smaller than previous reports for RP which suggested a figure of around 88hz. The jump down onto the accented syllable in Belfast rises is also something Rahilly (1991, p.111) reported from her data.[9] The $f0$ excursion differences between statements and questions also raise the issue of whether this potentially paralinguistic phenomenon is indicative of linguistic change or not.

4.5.3 Assessment

This initial analysis of IViE data from Belfast and Cambridge English in the parameters of alignment and $f0$ excursion found that overall there was less phonetic distance between Belfast statements and questions (from Belfast or Cam-

---

[9]The 'jump down' in pitch would fit with part of the chapter 3 theoretical proposals on intonational alignment change, but is over-ridden by the much clearer connection between Belfast statements and Belfast questions than between Belfast statements and Cambridge statements.

bridge English) than between Belfast statements and Cambridge statements. The large alignment difference between Belfast and Cambridge statements goes against the Alignment hypothesis (section 4.2.1). Therefore, it is difficult to reconcile the theoretical account of change from chapter 3 with these IViE data.

The Transfer hypothesis represents a more parsimonious account of the origin of Belfast statements, particularly the small degree of distance between Belfast statements and Belfast questions. Yet Belfast statements were significantly different to Cambridge questions in $f0$ excursion. There was also a trend for Belfast questions to have a wider $f0$ excursion than Belfast statements (see Ladd, 2008 for the similar suggestion in relation to Glasgow English). There does not appear to be a discrete intonational difference between Belfast English statements and questions. The parameter of $f0$ excursion emerges as gradiently variable in previous studies e.g. Ladd and Morton (1997).

The conclusion I draw from Experiment I is that I need to explore other options for the source of Belfast nuclear statement intonation than a phonetically modified nuclear statement intonation from another English variety or standard question intonation. I have shown an initial attempt to use the parameters of alignment and $f0$ excursion to assess intonational distance.

4.5.4 *Expansion of Transfer hypothesis*

Nuclear rises are not just typically associated with question intonation but also with Continuation intonation (Gussenhoven, 2004; Ambrazaitis, 2008), showing that the speaker is not finished their turn. Similarly, when a speaker is reciting a list of items, non-final items often also contain rising intonation in contrast to a fall on the final list item. Cruttenden (1997, p.96) also describes related functions of low rises as "non-finality, dependence, and syntactic subordination". Dombrowski and Niebuhr (2005) found specific phonetic differences between German rises that were 'turn-holding' in discourse (Continuation rises) and rises that were 'turn-yielding'. Though there was not a large difference in overall $f0$ excursion between these two types of German rises, there was a strong difference in the amount of $f0$ excursion of the rise on the accented syllable itself. Could Belfast English speakers have transferred continuation/list intonation to a general statement context?[10] Transfer of either question intonation or continuation intonation to a statement context would suggest that intonational change

---

[10]This possibility is based partly on explorations of archive data of Northern Irish English to which I refer again in chapter 5.

may not be intrinsically phonetic in nature but rather pragmatic/discourse motivated.

4.6 New experiment needed

The Alignment and Transfer hypotheses need to be examined in a larger-scale experiment with specifically designed materials. I designed Experiment II to address the need for the following important points:

- more appropriate data by which to examine the realisation of statement rises, question rises and continuation rises;
- directly comparable data from a few speakers of Belfast and Glasgow English.

The Alignment hypothesis could benefit from the choice of different English varieties against which to compare Belfast English. If there has been phonetic change between nuclear statement intonation like that found in Cambridge English and Belfast nuclear statement intonation, it might be that Cambridge and Belfast are on opposite ends of a cline. If there were intermediate varieties between them, intonational distance measurements might link up better with my theoretical account of phonetic change than in the initial IViE analysis. There is a precedent for this possibility in Hognestad (2006). He claims that the Stavanger accent I data show the crucial intermediate stage necessary to link West and East Norwegian varieties. The most obvious choice for me is to include Glasgow English. Belfast and Glasgow English have both been included in the UNB statement rises phenomenon but they have never been directly compared. Experiment II involved a few each of Belfast English speakers, Glasgow English speakers and additionally, Edinburgh English speakers. Edinburgh is closer geographically to Belfast and Glasgow than Cambridge but yet has not been traditionally included among the UNB varieties (Brown et al., 1980; Cruttenden, 1997; Aufterbeck, 2003).

In Experiment I, I could evaluate the Transfer hypothesis only in a very restricted way. The Cambridge question rises were heavily biased towards declarative questions in contrast to the Belfast questions. Importantly, I also only examined nuclear syllables with one following unstressed syllable. This may not be enough to identify key phonetic differences between statement and question intonation e.g. the difference between 'rise-plateau-(slump)' and continued rise (see section 4.5.1 above and section 4.12.5 below). Experiment II was specially

designed to deal with this shortcoming. It contained a set of words/compounds in nuclear position with three following unstressed syllables. In addition to a Sentence Reading task, participants in the new experiment also took part in semi-spontaneous Map Task games (Anderson et al., 1991; Lickley et al., 2005). This also allows a more natural discourse context for the statement vs. question distinction.

4.7 Experiment II

4.7.1 Introduction

This section deals with Experiment II, in which I did phonetic analysis of intonational nuclear tones in data from a few speakers of Belfast and Glasgow English.[11] I give a reminder of the hypotheses (section 4.7.2) and explain the methodology (section 4.8). Then I present the results of the Sentence reading task, addressing the implications for the hypotheses (section 4.9). I show the importance of Glasgow English data for assessing the Alignment hypothesis (section 4.11), and make detailed comparisons of statements, questions and list items for the evaluation of the Transfer hypothesis (section 4.12). I argue that the Belfast data supports the Transfer hypothesis, with transfer more likely from continuation/list rises than from question rises. I proceed to some analysis of more spontaneous Map task data (section 4.13) and show how this provides new support for my arguments about the nuclear rises from these particular Belfast and Glasgow English speakers. Finally, I include a short section on a few extra parameters that may help to evaluate the hypotheses (section 4.14).

Experiment II enables me to argue that there may be a different source for the rises in these particular Belfast and Glasgow English speakers respectively.

4.7.2 Hypotheses and parameters

The two hypotheses I propose in relation to the origin of nuclear statement intonation in Belfast and Glasgow English are:

1. the Alignment hypothesis
2. the Transfer hypothesis.

---

[11]and Edinburgh English

To evaluate these hypotheses, I make reference to intonational phonetic distance measurements along the following primary parameters:

1. Alignment of the $f0$ minimum at beginning of the rise (L1);
2. Alignment of the $f0$ maximum end of the rise (H1);
3. Scaling of L1;
4. Scaling of H1.

The Alignment hypothesis states that statement intonation in Belfast and Glasgow English arose through rightward "Re-alignment" of the nuclear peak (Dalton and Ní Chasaide, 2005; Dalton, 2007), transforming a nuclear fall into a nuclear rise. In contrast to Experiment I in which I made cross-variety comparisons, I evaluate the Alignment hypothesis mainly through within-variety comparisons of utterances with differing numbers of syllables following the nuclear syllable (section 4.11 below).

The Transfer hypothesis states that statement intonation in Belfast and Glasgow English involves the transfer of question intonation (Bolinger, 1978) or continuation/list intonation (Gussenhoven, 2004; Dombrowski and Niebuhr, 2005) to ordinary statements. Within-variety comparisons of statements, questions and lists, using measurements from the primary phonetic parameters is my approach to evaluating this hypothesis (section 4.12 below).

4.8 Methodology

4.8.1 *English varieties and speakers*

I recorded a few female speakers of Belfast, Glasgow and Edinburgh English in Autumn 2009. I chose speakers who were born in and had lived for most of their lives in the relevant city. Their parents/guardians also grew up in a similar area. The speakers considered that they had a typical accent of their respective variety. Their ages ranged from 18 to 25 and all were students currently living in Edinburgh. Though I did not elicit this information directly, the reader may assume that all of the Belfast and Glasgow English speakers with the exception of AG had lived in Edinburgh since the age of 18/19. AG (aged 23, see tables below) had only lived in Edinburgh for a couple of months at the time of the recording. One speaker (CC, Edinburgh) underwent speech therapy as a child though no other speakers reported any speech or hearing difficulties. I paid speakers a

small sum for their participation in both of my experimental tasks (£4 for the Sentence Reading task, £8 for the Map task). Further details on each individual speaker are given below.

I report analysis of data from three speakers in each variety:

1. EL, PT, RB (Belfast);
2. AG, CD, CR (Glasgow);
3. AR, CC, SK (Edinburgh).

I acknowledge that referring broadly to the varieties of Belfast, Glasgow and Edinburgh involves making large generalisations. I did not take sociolinguistic information into account in the choice of speakers or examine potential stylistic/register differences on any large scale. This proved to be a large shortcoming, particularly in relation to the Glasgow data. Unfortunately, only one of the Glasgow speakers consistently produced the nuclear rising phenomenon on statements (AG). The other two speakers did so to a lesser extent.[12] Once this problem emerged, I sought further speakers of Glasgow English but they too produced nuclear falls overwhelmingly on statements. I recorded seven Glasgow English speakers in total.

The problem of gathering enough data of Glasgow nuclear statement intonation is something the reader should bear in mind when assessing the results of the experiment. Still, this analysis provides one of the first ever direct comparisons of Belfast and Glasgow nuclear intonation. To provide some justification that the Glasgow nuclear statement contours I report in this chapter are representative of Glasgow nuclear statement intonation more generally, I consulted data from the HCRC Map task corpus of Glasgow English (Anderson et al., 1991).

4.8.2 *Design of test materials*

Speakers were recorded taking part in two tasks: a Sentence Reading task and a series of Map task games.

---

[12]Cruttenden (2007, p.271) considers that UNB intonation may "be more strongly established and more pervasive in Belfast than in Glasgow."

Initials	AG	CD	CR
Age	23	18	25
From	Lenzie, North Glasgow	South Glasgow	Northeast Glasgow
No. of years living there	23 years	18 years	17.5 years
Father from	North Glasgow	South Glasgow	Govan/Bearsden (in Glasgow)
Mother from	North Glasgow	South Glasgow	Buckie but lived in Glasgow for 35 years
Accent	typical	typical	fairly typical
Studying	Psychology	Law	PhD School of Community Health Sciences
Fluency in other languages	none	none	none
Tasks	Sentence reading and Map task	Sentence reading and Map task	Sentence reading only

Table 4.1: Details of the three female Glasgow English speakers whose data I use in Experiment II

Initials	EL	PT	RB
Age	22	22	20
From	Southeast Belfast	East Belfast	South Belfast
No. of years living there	19	18	18
Father from	Southeast Belfast	Northeast Belfast	East Belfast
Mother from	Southeast Belfast	Northeast Belfast	County Tyrone
Accent	typical	typical	typical
Studying	Biology	Speech and Language Therapy	English literature and Spanish
Fluency in other languages	none	none	none
Tasks	Sentence reading and Map task	Sentence reading and Map task	Sentence reading and Map task

Table 4.2: Details of the three female Belfast English speakers whose data I use in Experiment II

4.8.3   Sentence reading task

The Sentence reading task comprised 252 sentences in total, which were read aloud by participants. The task contained target words/compounds in sentence-final position embedded in four different sentence types:

1. statement,
2. a list of three clauses, of which the first clause was of interest (henceforth 'list'),
3. wh question,
4. y/n inversion question.

In choosing target words/compounds, I tried to maximise the presence of sonorant and voiced segments as this makes $f0$ analysis more straightforward. I expected participants to produce the nuclear pitch accent on the primary stressed syllable of the target words in this position in the sentence.

Here is an example of a target word/compound embedded in each sentence type:

1. They spent time in the **amusement arcade**.
2. They spent time in the **amusement arcade**, sped around the roller coaster and loved the bungee jumping.
3. When did they spend time in the **amusement arcade**?
4. Are you sure? Did they spend time in the **amusement arcade**?

Each y/n question was preceded by the question 'Are you sure?' This was intended to increase the questioning attitude. I believed this would increase the likelihood of speakers of Edinburgh English in particular, producing nuclear rises on these questions (Haan, 2002) as opposed to using other kinds of question intonation (Cruttenden, 1997).[13]

I included questions and lists to address the Transfer hypothesis. Dombrowski and Niebuhr (2005, p.192-3) observe that when working with the functional definitions of questions and continuation, this is not actually comparing like with like. Questions usually have a separate syntactic form to statements. Continuation refers more to the discourse function of incompleteness. The lists (my

---

[13]In fact, I refer only to data from Edinburgh English wh questions in this chapter as the inclusion of 'are you sure' was not as successful in eliciting nuclear rises in Edinburgh y/n questions as I had hoped.

instantiation of Continuation in the current study) and two question types in the Sentence Reading task for Experiment II are clearly defined, however.

I grouped the target words/compounds into two main conditions:

- 'Short tails',
- 'Long tails'.

In the 'Short tails' condition, the target words contained initial stress and just one following unstressed syllable (i.e. there was just one syllable in the 'tail' of the nuclear tone, using the terminology of the British tradition). There was a further sub-division within this condition between target nonsense words and target real words. All participants uniformly produced the nonsense words with initial stress and in general had no pronunciation difficulties with them. The reason I used a set of nonsense words was to design controlled materials to study spectral tilt and peak amplitude in nuclear rises (see section 4.14 below). These required the stressed syllable and following unstressed syllable to be segmentally identical e.g. 'lala'.

There were 6 target nonsense words and 11 target real words in the 'Short tails' condition. Each sentence with a nonsense target word was read five times by each participant. This led to a total of 120 utterances (6 target words x 4 sentence types x 5 renditions) per speaker in this sub-division of the 'Short tails' condition. Each sentence with a target real word was read once by each participant. This led to a total of 44 utterances per speaker (11 target words x 4 sentence types) in the second sub-division of this condition.

In the 'Long tails' condition, there were three or four unstressed syllables following the primary stress (i.e. there were several syllables in the 'tail'). I achieved this by using long compounds with primary stress on the initial or second syllable from the left e.g. 'amusement arcade', 'war memorial'. There were 22 target words in the 'Long tails' condition. Each target word was read once in each of the four different sentence types. Therefore, there were 88 sentences (22 x 4) per speaker in this condition.

The sentences were randomised and each participant read the sentences in a different randomised order. The sentences were printed onto A4 paper (double-sided) and in 20 pt Times New Roman font. The A4 pages were placed on a music stand. I appended five sentences each to the beginning and end of the

randomised set of sentences. I omitted these from any analysis. These extra sentences were repetitions of some of the test sentences so that participants would not realise that they formed a preface and conclusion to the main set of sentences. I did not include any filler sentences. This is because the diverse nature of the materials in the 'Short' and 'Long tails' conditions (nonsense words vs. real words, four different sentence types) resulted in sentences from one condition essentially acting as fillers for sentences in the other condition.

I told participants that the Sentence reading task included a mixture of statements and questions, that there would be some nonsense words and some repetitions of sentences. I asked them to read the sentences at a normal speaking rate. If they made a mistake, I told them that they could either repeat the sentence or continue straight on.

The analysis I present in this chapter is based on a subset of the complete Sentence reading task data. I include analysis of 40 utterances per speaker from the nonsense word sub-division of the 'Short tails' condition (2 target words x 4 sentence types x 5 renditions) and 40 utterances per speaker from the 'Long tails' condition (10 target words x 4 sentence types). This is for the Belfast and Glasgow English speakers. I also include analysis of 10 utterances per speaker from the Edinburgh English data (wh questions only). The relevant sentences from the 'Short' and 'Long tails' conditions are given in appendix D.

4.8.4 Map task

The Map task technique is a method for eliciting speech that is not as controlled as read speech but still allows the researcher more control over the data than fully spontaneous speech (Anderson et al., 1991). Lickley et al. (2005), McGregor (2005) and Warren (2005) are some of the most prominent previous studies to have used the technique for studying intonation. The IViE project (Grabe et al., 2001; Slater, 2008) also included the Map task but none of the resulting publications refer to any large-scale analysis of it.

The Map task is a game which participants play in pairs. One participant is the Instruction Giver and the other, the Instruction Follower. Both players have a specially designed map with a start point and landmarks on it. The Instructions Giver's map also contains a route drawn on it and a finish point. The Instruction Giver's task is to explain this route to the Instruction Follower so that the Instruction Follower may draw it on their map. The idea is that the landmarks on the

map should help the Instruction Giver to explain the route and the Instruction Follower to understand it. The names of the landmarks can be carefully chosen by the researcher to test their own hypotheses. A further aspect of the design to elicit renditions of the landmark names in a more natural conversation setting is that there are mismatches between the Instruction Giver's and Follower's maps. For example, there may be landmarks present on one player's map that are absent on the other player's map.

The landmark names I used were the same as many of the target real words in the Sentence reading task.[14] As in that task there was a distinction between 'Short tails' and 'Long tails'.

I designed 4 pairs of maps, which were printed on A3 paper in colour. Participants were seated at a table in separate rooms. There was a window between the two rooms enabling the participants to see each other but they were not positioned directly in front of this window. I told participants that the Instruction Giver's map had the route drawn on it and that the Instruction Giver had to explain the route to the Follower so that she could draw it on her map. I told them that the landmarks would help them to explain the route but that there were some mismatches between the Instruction Giver's and Follower's maps. I also told them that they were free to say anything that they wished during the game and that they could consider the game as an opportunity for free conversation.

I used the Quad design outlined in Anderson et al. (1991, p.360-1). Following this design, four participants from a given variety took turns in a special order to play the games involving all four map pairs. Each player was the Instruction Giver twice and the Instruction Follower twice. This led to eight conversations in total.

The main value of the Map task technique for my hypotheses is that it elicits several renditions of questions and non-questions. Given the more spontaneous nature of the speech in the Map task context, it would have been extremely difficult for me to attempt to make subtler distinctions within the two broad categories of question and non-question. Even the distinction between them was not always clear-cut. Non-questions included commands/instructions, responses to commands/instructions and explanations. Non-questions form a counterpart to statements in the Sentence reading task.

---

[14]I should clarify that I did not use any nonsense words in the Map task design.

Variety	'Tail'	Number of IPs
Belfast	Short	44
Belfast	Long	45
Glasgow	Short	21
Glasgow	Long	23

Table 4.3: Breakdown of Map task utterances by Variety and 'Tail'

Variety	Utterance type	Number of IPs
Belfast	question	22
Belfast	non-question	67
Glasgow	question	21
Glasgow	non-question	23

Table 4.4: Breakdown on Map task utterances by Variety and Utterance type

Within-variety comparison of statements in the 'Short' and 'Long tails' conditions allows another context in which to address the Alignment hypothesis (see section 4.7.2 above). The comparison of questions and non-questions within each variety is relevant to the Transfer hypothesis.

I focus on data from three of the four Belfast English speakers who took part in the Belfast Map task games (EL, PT, RB) and two of the four Glasgow English speakers who took part in the Glasgow Map task games (AG, CD).

I refer to analysis of 89 Belfast English utterances and 44 Glasgow English utterances.

There is a further breakdown of these utterances by speaker in appendix D.

4.8.5 Recording procedure

All recordings took place at the joint PPLS[15]/Informatics recording facility at the University of Edinburgh. All speakers wore DPA 4035 head set condenser microphones. The acoustic signal was recorded on .wav files in mono for the Sentence reading task and stereo for the Map task. It was digitised at a 48kHz sampling rate/16 bit.

---

[15]School of Philosophy, Psychology and Language Sciences

4.8.6 File analysis

The Praat scripts and segmentation criteria that I used for the analysis of Experiment II included the same scripts as I used in Experiment I. I modified these to make them appropriate for the different test materials in the new experiment. For example, in the 'Long tails' condition, I segmented the three following unstressed syllables after the nuclear stress.[16]

For the alignment measurements, the script extracted the $f0$ maximum (H1) in the target region of the nuclear syllable, preceding syllable and 'Tail' syllables. It also extracted the $f0$ minimum (L1) preceding H1.

For the $f0$ excursion and scaling measurements, the relevant values extracted by the script in each utterance were:

1. the $f0$ value of L1 and H1 in Hz (linear) and logHz;
2. the $f0$ excursion between L1 and H1;
3. the mean $f0$ value of the utterance in Hz (linear) and logHz scales;
4. the duration of the utterance in milliseconds (ms).

I did not get a measure of the overall speech rate of each utterance in Experiment II. However, the script extracted the durations of the syllables in the nuclear tone and a measure of the rate of the nucleus and 'tail' stretch (total duration of the nuclear syllable and 'tail' divided by the number of syllables in the nuclear syllable and 'tail').

4.8.7 Unusable data

I applied the following criteria to the Sentence reading task data and Map task data for discarding files from further analysis:

1. files containing nuclear downstepped falls;
2. files with major $f0$ perturbations or tracking errors which prevented reliable locations of $f0$ minima and maxima from being found in the target region;
3. files with clear hesitations and disfluencies;
4. files with extremely flat intonation;

---

[16]'War memorial' was the only target compound word in the 'Long tails' condition with four following unstressed syllables. However, I segmented this as 'me.mo.rial' since the final vowel was so reduced.

5. files in which the target word/compound had been deaccented.

In the Sentence reading task, 80 sentences spoken by three Belfast English speakers and three Glasgow English speakers would lead to a total of 480 utterances for analysis. Due to the above problems, I had to discard 106 of these utterances from further analysis, leaving 374 utterances. 76 of these 106 utterances were discarded due to downstep, flat intonation or an otherwise inappropriate contour. The vast majority of these were from the Glasgow English speakers. 18 utterances were discarded due to deaccenting, 9 due to major perturbations/errors and 3 due to major hesitations.

4.9 Results: Sentence reading task

4.9.1 Statistical analysis

The statistical analysis I used for the data from Experiment II involved Linear Mixed-Effects Models (Baayen, 2008; Baayen et al., 2008). I conducted the analysis using R and the special LanguageR library of packages. I used the **lmer** (Linear Mixed Effect Regression) function. Mixed-Effects models allow the researcher to model their data using fixed and random effects. Random effects are those which the researcher believes may randomly affect their data. Speaker and items/words are two common such factors. Fixed effects are factors which the researcher believes will not randomly affect their data. These factors may also be measured again at the same levels in the future; random effects may not (Bates, 2005; Mills, 2008). For example, the factor Variety with levels of Belfast, Glasgow and Edinburgh English could be measured again in a subsequent study. However, a new study of Belfast and Glasgow English would use different participants to the ones I used in my study. Baayen et al. (2008, p.405) and Baayen (2008, p.289-290) report that Mixed-Effects models are robust to unequal sample sizes and to missing data, both of which are relevant to my data.

To assess p values, I used Markov Chain Monte Carlo (MCMC) sampling with 10,000 samples (**pvals.fnc** command in the LanguageR library). P values from MCMC sampling do not suffer the problem of an anti-conservative bias that the t-distribution has for small sample sizes (Baayen et al., 2008, p.398). In addition, MCMC sampling does not have underlying assumptions that the data are

normally distributed or have equal variance (Mills, 2008, p.62ff.). Such assumptions do underlie other methods of analysis, such as Repeated Measures ANOVA (Quené and van den Bergh, 2004, 2008).

I used the fixed effects of Variety (2 levels: Belfast, Glasgow)/(3 levels: Belfast, Glasgow, Edinburgh), 'Tail' (2 levels: Short, Long) and Sentence type (4 levels: Statement, list, wh question, y/n question). The random effects were speaker (6 levels)/(9 levels) and word (12 levels). The main dependent variables were the phonetic parameters outlined in section 4.7.2 above.

In the following sections, I present only the statistical results that are directly relevant to the evaluation of the Alignment and Transfer hypotheses. Further tables of statistical analysis are included in appendix D.

4.9.2   L1 alignment

I report alignment results as measured from the beginning of the nuclear vowel (Begv1), expressed as a proportion of the duration of v1. The reason for this choice of landmark was my observation that in the Glasgow data (statements and wh questions in particular), L1 tended to be aligned near the beginning of v1. The alignment of L1 in the Belfast data appeared to be later than this, more towards the end of v1, similar to my findings from the IViE Belfast data in Experiment I. I present the mean alignment values here in Tables 4.3 and 4.4.

A minus value indicates that L1 occurred before the beginning of v1. Mayo (1996) and Mayo et al. (1997) did not identify a particularly strong alignment location for the $f0$ minimum at the beginning of Glasgow rises. They do report a tendency for the $f0$ minimum to be aligned before the beginning of the accented syllable. This tendency is only paralleled here in the 'Short tails' statements and wh questions.

The mean values for the Belfast data indicate that L1 is aligned well beyond the beginning of v1. A value of 0.886 (Belfast 'Long tails' statements), for example, indicates that the distance between Begv1 and L1 takes up nearly 90% of the duration of v1. A value of -0.033 (Glasgow 'Long tails' wh questions) indicates firstly that L1 was aligned before Begv1 and secondly, that the distance between L1 and Begv1 corresponds to c.3% of the duration of v1.

'Short tails' Sentence type	Belfast	Glasgow
statement	0.573 (s.d. 0.42)	-0.471 (s.d. 0.651)
list	0.689 (s.d. 0.326)	0.603 (s.d. 0.838)
wh question	0.382 (s.d. 1.496)	-0.476 (s.d. 0.662)
y/n question	0.45 (s.d. 0.397)	0.156 (s.d. 1.063)

Table 4.5: Mean proportional alignment values for L1 from Begv1 ('Short tails')

'Long tails' Sentence type	Belfast	Glasgow
statement	0.886 (s.d. 0.587)	0.228 (s.d. 0.788)
list	1.3 (s.d. 1.093)	5.072 (s.d. 8.641)
wh question	0.414 (s.d. 0.899)	-0.033 (s.d. 1.327)
y/n question	0.525 (s.d. 0.88)	1.726 (s.d. 2.642)

Table 4.6: Mean proportional alignment values for L1 from Begv1 ('Long tails')

For quasi-comparability with the IViE data from Experiment I, I also measured the alignment of L1 as the absolute distance from the end of v1. The relevant details are in appendix D. They indicate that L1 in Experiment II was earlier aligned than L from Experiment I. Direct comparisons between the IViE Belfast data and the Belfast data from Experiment II are not appropriate because L at the beginning of the rise was measured as the 'elbow' in the IViE data, but L1 in the present experiment was the true $f0$ minimum. If the 'elbow' marks the point of greatest acceleration of the rise, then the true $f0$ minimum would be aligned earlier than this anyway though.

4.9.3 H1 alignment

I report the alignment of H1 from Endv1 as a proportion of the duration of the 'tail' syllables. I chose this segmental landmark because the alignment of H1 in the Glasgow 'Short tails' appeared to be close to the end of the nuclear vowel. The alignment of H1 in the Glasgow 'Long tails' and in general in the Belfast data appeared to be later than this. I present the mean alignment values in Tables 4.5 and 4.6.

'Short tails' Sentence type	Belfast	Glasgow
statement	0.82 (s.d. 0.138)	-0.182 (s.d. 0.596)
list	0.839 (s.d. 0.189)	0.732 (s.d. 0.494)
wh question	0.726 (s.d. 0.144)	0.039 (s.d. 0.478)
y/n question	0.629 (s.d. 0.121)	0.38 (s.d. 0.531)

Table 4.7: Mean proportional alignment values for H1 from Endv1 ('Short tails')

'Long tails' Sentence type	Belfast	Glasgow
statement	0.84 (s.d. 0.184)	0.235 (s.d. 0.215)
list	0.909 (s.d. 0.174)	0.651 (s.d. 0.426)
wh question	0.741 (s.d. (0.19)	0.332 (s.d. 0.304)
y/n question	0.723 (s.d. 0.211)	0.541 (s.d. 0.392)

Table 4.8: Mean proportional alignment values for H1 from Endv1 ('Long tails')

A proportional alignment value of 0.82, for example, represents that the distance between Endv1 and H1 takes up just over 80% of the duration of the 'tail'. A proportional alignment value of -0.182 represents first of all that the distance between Endv1 and H1 takes up 18% of the duration of the tail. The minus value indicates further that H1 actually occurs before the tail.

4.9.4 'Nuclear tone rate'

In three languages, van Heuven and van Zanten (2005) found a faster speech rate (measured in syllables per second) in questions than in statements. In such a case, it is possible that subtle alignment differences between statements and questions in an individual variety might be an artifact of speech rate or segmental durational differences between the sentence types. Incorporating 'Nuclear tone rate' as a random factor would allow me to examine if there were still effects of alignment differences beyond rate differences. It did not make a difference to the results, indicating that the alignment differences I report below (sections 4.11 and 4.12) are valid.

However, I also examined specifically whether there were 'Nuclear tone rate' differences in the data.

'Short tails' Sentence type	Belfast	Glasgow
statement	0.229 (s.d. 0.015)	0.22 (s.d. 0.012)
list	0.231 (s.d. 0.021)	0.228 (s.d. 0.021)
wh question	0.225 (s.d. 0.018)	0.216 (s.d. 0.022)
y/n question	0.22 (s.d. 0.016)	0.207 (s.d. 0.018)

Table 4.9: Mean 'nuclear tone rate' (duration divided by no. of sylls) ('Short tails')

There were some differences between sentence types. Y/n questions had a faster 'Nuclear tone rate' than statements ($pmcmc < 0.01$).[17] This is in line with van Heuven and van Zanten (2005) although I acknowledge that 'Nuclear tone rate'

---

[17]The p value was derived by Markov Chain Monte Carlo (MCMC) sampling, which is why I use 'pmcmc'.

'Long tails' Sentence type	Belfast	Glasgow
statement	0.177 (s.d. 0.028)	0.182 (s.d. 0.031)
list	0.189 (s.d. 0.027)	0.193 (s.d. 0.031)
wh question	0.176 (s.d. 0.026)	0.177 (s.d. 0.022)
y/n question	0.166 (s.d. 0.016)	0.142 (s.d. 0.018)

Table 4.10: Mean 'nuclear tone rate' (duration divided by no. of sylls) ('Long tails')

and speech rate are not directly comparable. Further, y/n questions had a faster 'nuclear tone rate' than wh questions ($pmcmc < 0.01$). These rate differences, inasmuch as they are linked to syllable duration differences, would suggest that proportional alignment measurements may be more robust than absolute measurements. This is because the former take into account the duration of a segment/syllable and express the alignment of an $f0$ minimum or maximum as a proportion of this duration. For example, the $f0$ minimum (L1) may be aligned at roughly 25% of the proportion of the nuclear vowel in both statements and questions in a given variety, even if the nuclear syllable duration is shorter in the questions than in the statements. Absolute alignment measurements from the beginning of the nuclear vowel, let's say, would show the alignment of L1 as consistently earlier in the questions than in the statements simply because 25% of the nuclear syllable duration in the questions is a shorter absolute distance than 25% of the nuclear syllable duration in the statements.

4.9.5   Scaling normalisation

To try to get a somewhat speaker-independent assessment of potential differences between statements and questions in $f0$ height, I computed normalised values of the scaling of L1 and H1 separately. This has some connections with normalisation procedures for vocal tract length and vowel space used in acoustic segmental studies (see chapter 2).

Previous studies (including Patterson, 2000) suggest that speakers scale tonal targets "at the same level on their own speaker-specific scale...." (Ladd, 2008, p.199). To normalise for individual differences in overall speaker pitch range, I used the following procedure: For each speaker, I took all the individual utterance mean $f0$ values (logHertz) and calculated a weighted average based on the duration of each utterance. The formula is in appendix D. I subtracted the speaker's weighted average from the logHz value for L1 in each of that speaker's utterances. This gives a value for L1 normalised by the speaker's mean $f0$ (weighted

by utterance duration). It expresses how much below the speaker's mean $f0$ L1 occurred. I also subtracted the speaker's weighted average from the logHz value for each token of H1 for that speaker. This gives a normalised value for H1. This expresses how much above the speaker's mean $f0$ H1 occurred. I then combined the normalised values for L1 and H1 from different speakers when I conducted further statistical analyses. For clarity of presentation in this thesis, I express all logHz values as semitones. The formula for converting logHz to semitones is in appendix D. Nolan (2003) argues that the semitone scale (logarithmic) offers the best approximation of perceived pitch. Patterson (2000) also found in favour of the semitone scale. However, there is still much debate about the most appropriate scale to capture perceived pitch and indeed about the best way to normalise $f0$. Other studies have argued for the Equivalent Rectangular Bandwidth (ERB) scale (midway between linear and logarithmic below 500 Hz) (Glasberg and Moore, 1990; Hermes and van Gestel, 1991; Arvaniti et al., 2006; Arvaniti and Ladd, 2009). Early work by Rietveld and Gussenhoven (1985) favoured the linear Hz scale over the semitone scale, but their stimuli data came from a single female speaker and exploited only a restricted pitch register (Hermes and van Gestel, 1991). The semitone scale is suitable for use in this thesis, as it enables a clear depiction of differences in H1 scaling between statements and questions (see sections 4.9.7 and 4.12.4 below). The semitone values show how many semitones below or above the speaker's mean L1 or H1 occurred respectively. See also Fant and Kruckenberg (2004) for normalisation using the speaker's mean $f0$ expressed in semitones.

Patterson (2000, p.35) points out that pitch tracking errors may affect values like mean $f0$. I acknowledge that this may have been a problem in my data, though I took the mean $f0$ from the smoothed pitch object in Praat, which removed many of the pitch tracking errors.

4.9.6   L1 scaling

I present the mean normalised L1 scaling values for the Belfast and Glasgow data in each sentence type in Tables 4.9 and 4.10. The minus values indicate that these normalised L1 scaling values were below the speaker's mean $f0$.

'Short tails' Sentence type	Belfast	Glasgow
statement	-2.778 (s.d. 1.051)	-1.649 (s.d. 1.272)
list	-2.568 (s.d. 1.381)	-2.568 (s.d. 1.368)
wh question	-2.719 (s.d. 1.202)	-1.917 (s.d. 2.426)
y/n question	-2.352 (s.d. 1.458)	-2.434 (s.d. 1.211)

Table 4.11: Mean normalised scaling of L1 in semitones ('Short tails')

Long tails Sentence type	Belfast	Glasgow
statement	-2.274 (s.d. 1.188)	-1.202 (s.d. 0.96)
list	-1.231 (s.d. 0.757)	-1.694 (s.d. 1)
wh question	-2.394 (s.d. 1.372)	-1.364 (s.d. 1.385)
y/n question	-1.692 (s.d. 1.189)	-1.688 (s.d. 1.415)

Table 4.12: Mean normalised scaling of L1 in semitones ('Long tails')

4.9.7 H1 scaling

I present the mean normalised H1 scaling values for the Belfast and Glasgow data in each sentence type in tables 4.11 and 4.12. These are positive values indicating that H1 was scaled above the speaker's mean $f0$.

'Short tails' Sentence type	Belfast	Glasgow
statement	0.964 (s.d. 0.723)	0.136 (s.d. 1.161)
list	1.015 (s.d. 2.075)	0.016 (s.d. 0.899)
wh question	2.471 (s.d. 0.998)	1.991 (s.d. 1.412)
y/n question	4.141 (s.d. 1.391)	1.124 (s.d. 1.697)

Table 4.13: Mean normalised scaling of H1 in semitones ('Short tails')

'Long tails' Sentence type	Belfast	Glasgow
statement	1.808 (s.d. 2.718)	1.242 (s.d. 1.46)
list	2.369 (s.d. 1.14)	0.355 (s.d. 1.437)
wh question	4.551 (s.d. 1.075)	2.188 (s.d. 1.72)
y/n question	5.088 (s.d. 1.346)	2.514 (s.d. 1.879)

Table 4.14: Mean normalised scaling of H1 in semitones ('Long tails')

4.10   Implications

Now I turn to evaluating the implications of the alignment and scaling results for the two hypotheses:

1. the Alignment hypothesis
2. the Transfer hypothesis.

4.11 Alignment hypothesis discussion

There are two primary ways of addressing this hypothesis. The first is to compare the alignment of L1 and H1 between the varieties. The second is to compare the alignment of L1 and H1 between the 'Short tails' and 'Long tails' sets within each variety.

Comparing the alignment of L1 and H1 across the three varieties of Belfast, Glasgow and Edinburgh English would allow me to assess intonational distances between them (see section 4.6 above). I expected that Glasgow English might be intermediate between Edinburgh and Belfast English. Unfortunately, the majority of the Edinburgh English statements contained downstep and were thus unusable for comparison with the Glasgow and Belfast data. Nevertheless the within- and between-variety comparisons of these Belfast and Glasgow English speakers are in themselves important.

In within-variety comparison, a later alignment of H1 in the 'Long tails' than in the 'Short tails' would support the Alignment hypothesis. There are two plausible motivations for this. First, peaks (H1) have a well-known tendency to drift rightwards with an increase in the number of following unstressed syllables after a pitch-accented syllable. This synchronic allophony may lead to diachronic change (see chapter 3). If H1 in the 'Long tails' set (3 or 4 following unstressed syllables) shows evidence of being later aligned than H1 in the 'Short tails' set (one following unstressed syllable) in an individual variety, this may support the hypothesis that nuclear rises in this variety have come about through "Realignment" of the H beyond the nuclear syllable. Second, the Alignment hypothesis posits that nuclear statement intonation in Belfast and Glasgow English was originally realised with a nuclear fall more typical of statements in other varieties of English. This suggests the importance of a final low at the intonational boundary. If the realisation of a final low is obligatory, then later alignment of H1 in the 'Long' than the 'Short tails' would also be unsurprising. Given that there are only two syllables in the 'Short tails' in which to realise the nuclear tone (stressed and one unstressed syllable), H1 may need to be earlier aligned to avoid tonal crowding with the final low. By contrast, the greater number of syllables for nuclear tone realisation in the 'Long tails' may not result in tonal crowding between H1 and the final low.

It is less clear from previous work whether the initial L1 (preceding H1) would be expected to drift to the right alongside an increase in the number of following

unstressed syllables from the 'Short' to the 'Long tails'. L1 is arguably too far from a final low for issues of crowding to be relevant here. L1 appeared to be aligned more to the right in both varieties in the 'Long tails' than in the 'Short tails'. However, there was no statistically significant difference between the varieties or between 'Short' and 'Long tails' in the alignment of L1. Therefore, there are no clear-cut indications of "Re-alignment" of L1 in these data. Using a much larger data sample, Arvaniti et al. (2006, p.680) did find a difference in the alignment of the low elbow in Greek y/n question rise-falls based on the proximity of the nuclear syllable to the end of the IP.

4.11.1 Between-variety comparison

I proceed to discuss the alignment of H1. Here, there was a significant main effect of Variety i.e. a difference between Belfast and Glasgow ($pmcmc < 0.01$). Specifically, H1 was aligned proportionally much earlier in the Glasgow data than in the Belfast data. H1 in these Belfast data was usually aligned in the final syllable in the IP. All but eight utterances had H1 aligned in the final syllable. Syllable-final alignment was also the norm for the IViE Belfast data reported in Experiment I. This contrasts strongly with the Glasgow data, in which H1 was aligned much earlier. In the 'Short tails', it was aligned within the nuclear syllable and in the 'Long tails', was often aligned at least one syllable to the right of the nuclear syllable. The alignment difference in H1 between Glasgow and Belfast is a crucial finding in addressing the previous neglect of direct comparisons between these varieties. The reader should bear in mind that this conclusion is drawn from a very small sample of speakers. Further work is needed to ascertain if this alignment difference would also emerge among a much larger sample of Glasgow and Belfast English speakers.

4.11.2 Final vowel scaling

The other fundamental difference between these Glasgow and Belfast data relates to the pitch at the IP boundary. The former generally contain low pitch at the IP boundary whereas the latter contain high pitch. To reflect this difference quantitatively, I extracted the scaling of the beginning of the final vowel where the $f0$ was a non-spurious value and normalised this using the same procedure as in section 4.9.5 above. I expect there to be a difference between Belfast and Glasgow because at this time point, the Glasgow English contour is falling towards the final L, while the Belfast contour is rising towards the final H. My primary

Alignment H1: Statements
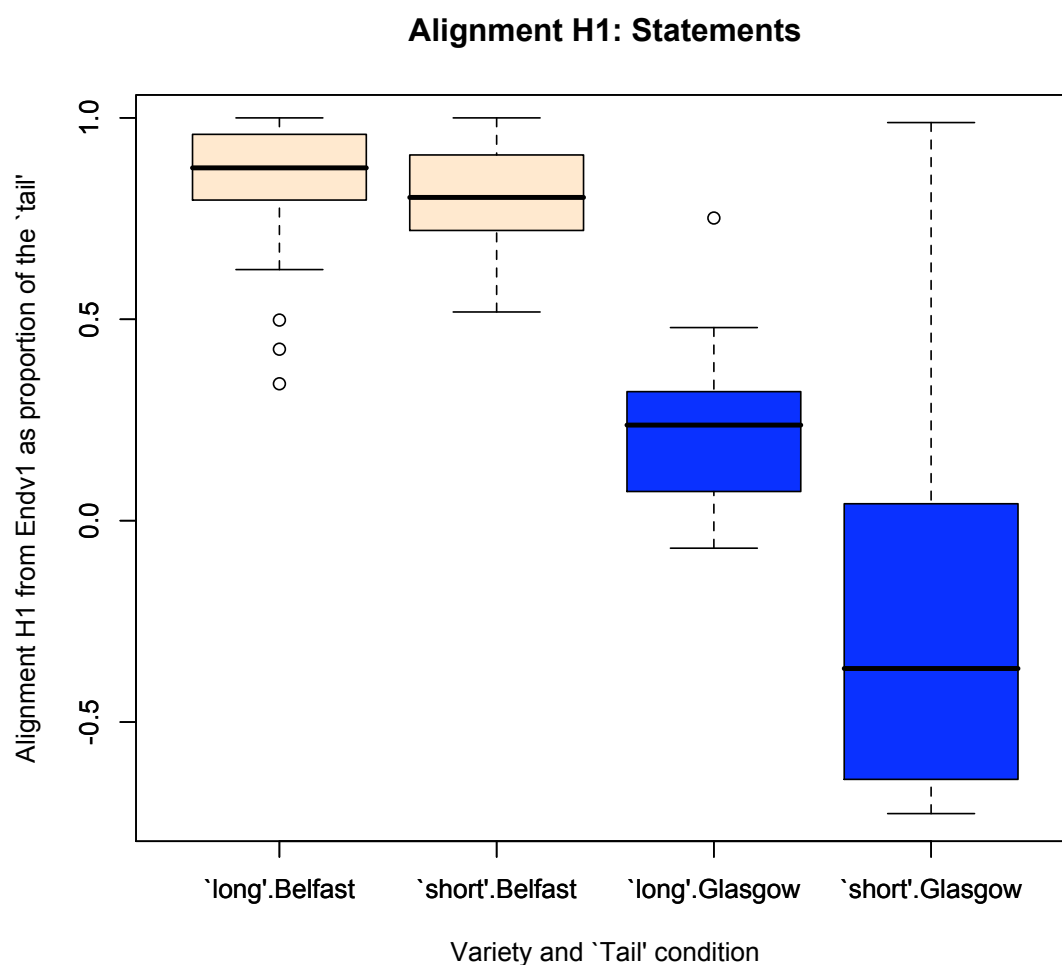
Figure 4.7: Boxplot displaying the alignment of H1 in 'Short' and 'Long tails' in Belfast and Glasgow English. For clarity, I have displayed the statement data only. Notice that H1, measured from Endv1, is aligned proportionally much later in both sets of Belfast statements than in the Glasgow statements. Notice also crucially the difference in alignment between the 'Short' and 'Long tails' in the Glasgow statements (see section 4.11.4 below).

interest was in the realisation of the nuclear rising element of the contour. This is why I did not have a specific measure of the alignment and scaling of the final low in Glasgow. The presence of a final low in the Glasgow data goes against Cruttenden's previous description of the typical Glasgow contour being a 'rise' rather than a 'rise-plateau-slump' or a 'rise-fall' (Cruttenden, 1997, p.133ff.). It supports Ladd (2008, p.127) who describes a "distinct fall" usually following the UNB rise. Again I remind the reader of the problems of gathering enough Glasgow data so further work is needed to be certain of representativeness.
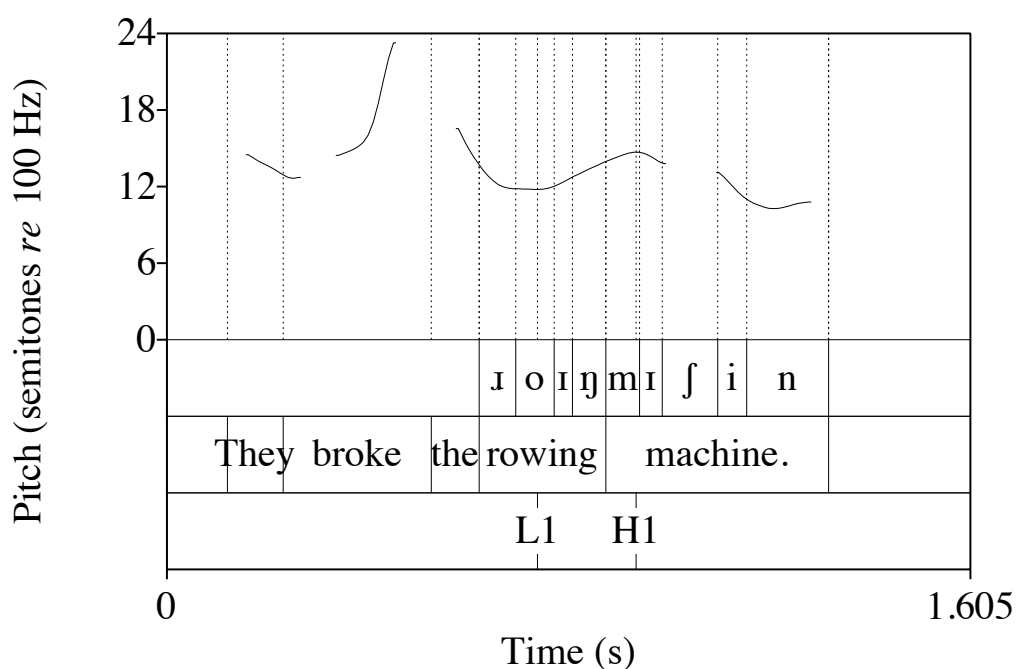
Figure 4.8: *f*0 trace of the statement 'She broke the rowing machine' spoken by Glasgow English speaker AG. Notice the alignment of H1 in the unstressed first syllable of 'machine.'

'Short tails' Sentence type	Belfast	Glasgow
statement	-1.534 (s.d. 1.332)	-1.96 (s.d. 1.367)
list	-1.552 (s.d. 1.454)	-2.383 (s.d. 1.273)
wh question	-0.876 (s.d. 1.956)	-0.948 (s.d. 1.856)
y/n question	0.903 (s.d. 2.706)	-1.738 (s.d. 1.745)

Table 4.15: Mean normalised scaling of onset of final vowel in semitones ('Short tails')

'Long tails' Sentence type	Belfast	Glasgow
statement	1.574 (s.d. 1.672)	-0.159 (s.d. 1.112)
list	0.623 (s.d. 1.194)	-1.121 (s.d. 0.767)
wh question	3.182 (s.d. 2.521)	-1.877 (s.d. 2.787)
y/n question	1.992 (s.d. 4.313)	-0.546 (s.d. 1.361)

Table 4.16: Mean normalised scaling of onset of final vowel in semitones ('Long tails')
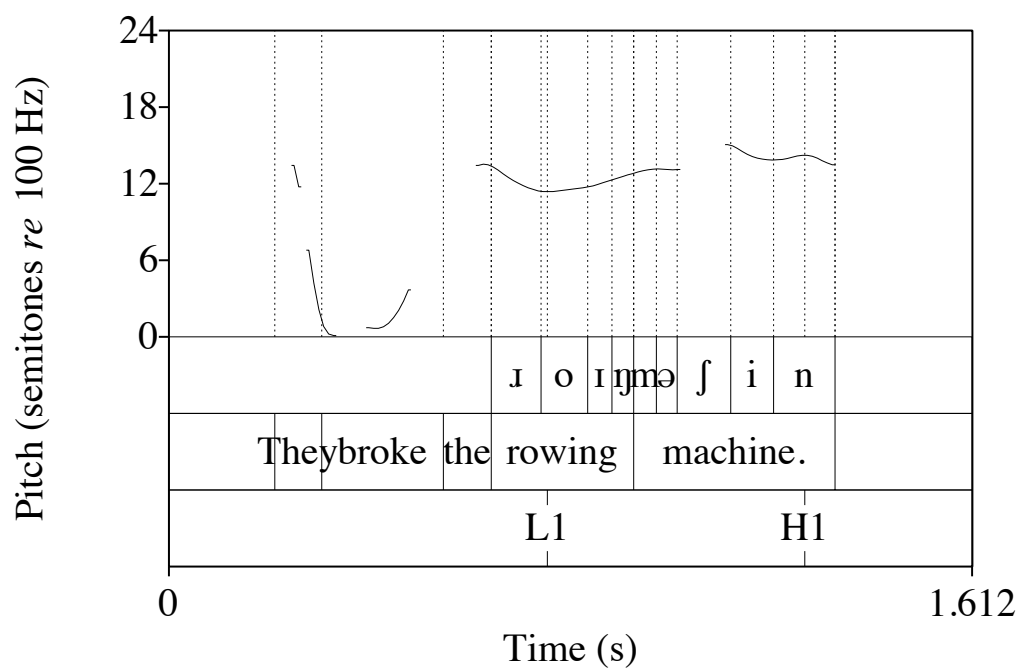
Figure 4.9: ʄ0 trace of the statement 'She broke the rowing machine' spoken by Belfast English speaker PT. Notice the alignment of H1 in the IP-final syllable.

Scaling of the beginning of the IP-final vowel: Statements
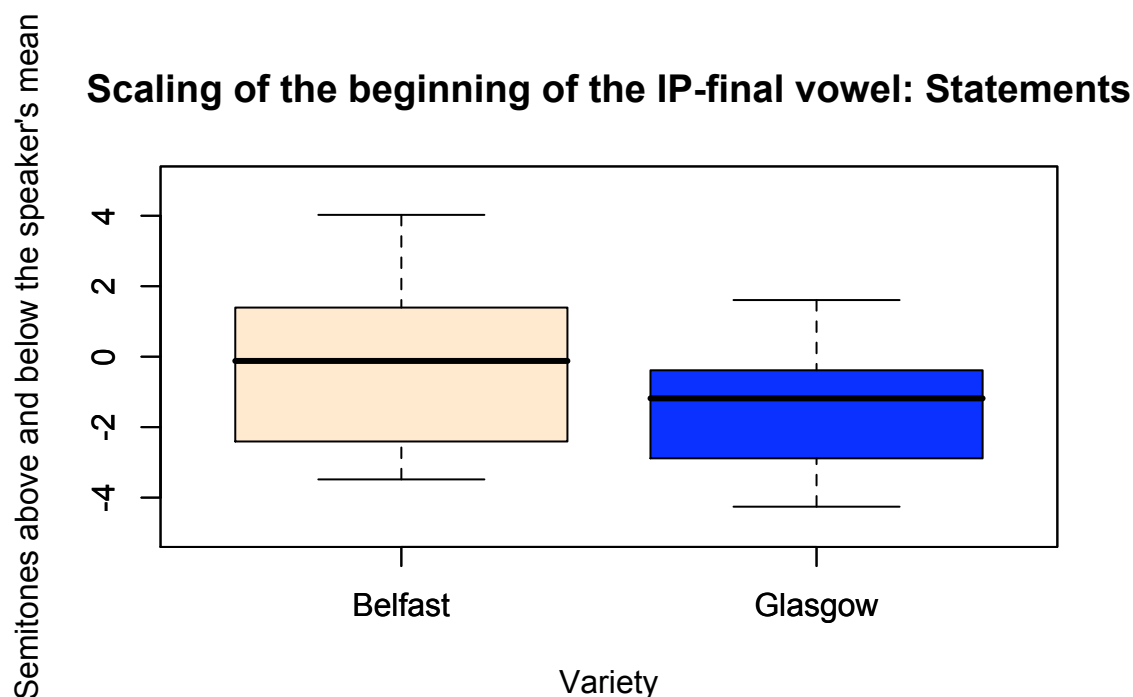
Figure 4.10: Boxplot displaying the scaling (normalised by the speaker's mean *f*0) at the onset of the IP-final vowel. For clarity, I have included the statement data only. Notice that the scaling is generally higher in the Belfast statements than in the Glasgow statements.

There was a significant main effect of Variety ($pmcmc < 0.05$). The mean scaling values (figures 4.13 and 4.14) show that the *f*0 at the onset of the IP final vowel is higher in the Belfast data than in the Glasgow data. For example, in the 'Long tails', Belfast statements are scaled 1.574 semitones above the speaker's mean *f*0 at the IP-final vowel onset, whereas the Glasgow statements are scaled 0.159 semitones below the speaker's mean *f*0.

There also appears to be some difference between the 'Long tails' and 'Short tails' generally. In the Belfast data, the pitch of the onset of the final vowel should be still quite low in the 'Short tails' whereas it should be quite high at this point in the 'Long tails'. The reasons for this are as follows:

In the 'Short tails' the final vowel is adjacent to the nuclear vowel. L1 in the Belfast data was consistently aligned within the nuclear vowel so at the onset of the subsequent vowel, the *f*0 is still fairly low (lower than speaker's mean *f*0). In the 'Long tails' the final vowel is not adjacent to the nuclear vowel so by the onset of the final vowel, the pitch has risen considerably from L1 (higher than speaker's mean *f*0). There is indeed a trend for the Belfast 'Long tails' to be

scaled higher than the 'Short tails'. This explains why the median scaling value of the combined 'Short' and 'Long tails' Belfast data are around the speaker's mean $f0$ (figure 4.10 above). In the Glasgow data, it might seem as if the 'Short tails' should have higher pitch at the beginning of the final vowel than the 'Long tails'. This was not borne out in the data, however. In the 'Short tails', there was a sharp fall over the onset consonant in the second syllable of the nonsense words 'lala' and 'vava'. So the pitch had already fallen considerably by the time point of the beginning of the final vowel.

The H1 alignment difference between Glasgow and Belfast, combined with the presence of the final low in Glasgow, leads me to consider the nuclear statement contours of the small set of speakers from the two varieties as phonologically distinct. The alignment of H1 in the Glasgow data strongly suggests phonological association to the nuclear syllable whereas the alignment of H1 in the Belfast data clearly supports a phonological association with the IP-boundary. The final low in the Glasgow data would fit with an L rather than a H tone associated to the IP-boundary.

The conclusion I have drawn here might make the reader wonder whether I have been faithful in my efforts to compare like with like in these intonational comparisons. If H1 in these Glasgow and Belfast data has different respective phonological associations, then perhaps I should not have compared them against each other in the first place. My response is as follows: the decision to compare the Glasgow H1 against the Belfast H1 was based on the Alignment hypothesis that historically they represented the 'same' element, which had drifted rightwards from the nuclear syllable. Internal comparisons of the Glasgow and Belfast data separately led to my conclusion that these Glasgow data supported the Alignment hypothesis whereas the Belfast data did not. Previous descriptions of these English varieties describe both as displaying a statement rising phenomenon so it was natural to begin with the assumption that the beginning and end points of those rises would be comparable between the varieties. This shows that one of my research questions was whether indeed the Glasgow H1 was strictly comparable to the Belfast H1. By contrast, studies of segmental phonetic distance do not seek to confirm whether one particular segment may be compared with another. Through knowledge of cognates and sound changes, they may already have a clear concept of comparing like with like and may proceed to other research questions. This is an important difference between my approach to phonetic

distances in intonation and how phonetic distances have been used in segmental studies.

4.11.3 *Preliminary resynthesis*

Separately, I tentatively experimented with resynthesising the $f0$ contour in Glasgow statements by moving the alignment of H1 to the right. I also experimented with moving the alignment of H1 of Belfast nuclear rises to the left. Though I did not conduct any formal analysis on this resynthesis, it indicated that a "Realignment" of H1 would not convert the Glasgow rise-fall into the Belfast rise. Dalton (2007) and Dalton and Ní Chasaide (2005) also demonstrated that "Realignment" of H1 in Cois Fharraige Irish would not produce the Donegal Irish contour. This further suggests the phonological distinctness of the Belfast and Glasgow nuclear statement contours.[18]

4.11.4 *Within-variety comparison*

This important difference between Belfast and Glasgow is furthered by within-variety differences in the behaviour of H1 alignment between 'Short' and 'Long tails'. Crucially, there was a significant interaction between Variety and 'Tail' ($pmcmc < 0.01$). There was a much greater difference in the alignment of H1 between the 'Short' and 'Long tails' in Glasgow than in Belfast.

In the Glasgow data, the alignment of H1 appears to show support for "Realignment" between 'Short' and 'Long' tails. There is evidence that the alignment of H1 is later with respect to Endv1 in the 'Long tails' set than in the 'Short tails' set. The mean proportional alignment value of -0.182 for the 'Short tails' statements shows that H1 was aligned towards the end of the nuclear syllable (within the last 18% of the nuclear vowel) (see section 4.9.3 above). This shows that in the 'Short tails', H1 was aligned before Endv1 and therefore within the nuclear syllable.[19] Dalton and Ní Chasaide (2005) suggest that an already late aligned peak within the accented syllable might be predisposed to drift rightwards beyond that syllable (see chapter 3).

---

[18]Even though the theoretical account of alignment change presented in chapter 3 allows for alignment change to lead to further phonetic changes, the Belfast data clearly do not show evidence of synchronic H1 alignment allophonic variability in the way that the Glasgow data do (section 4.11.4).

[19]I assume syllabification of the target nonsense words as 'la.la' and 'va.va'. Treatment of ambisyllabicity in English is far beyond the scope of this thesis.
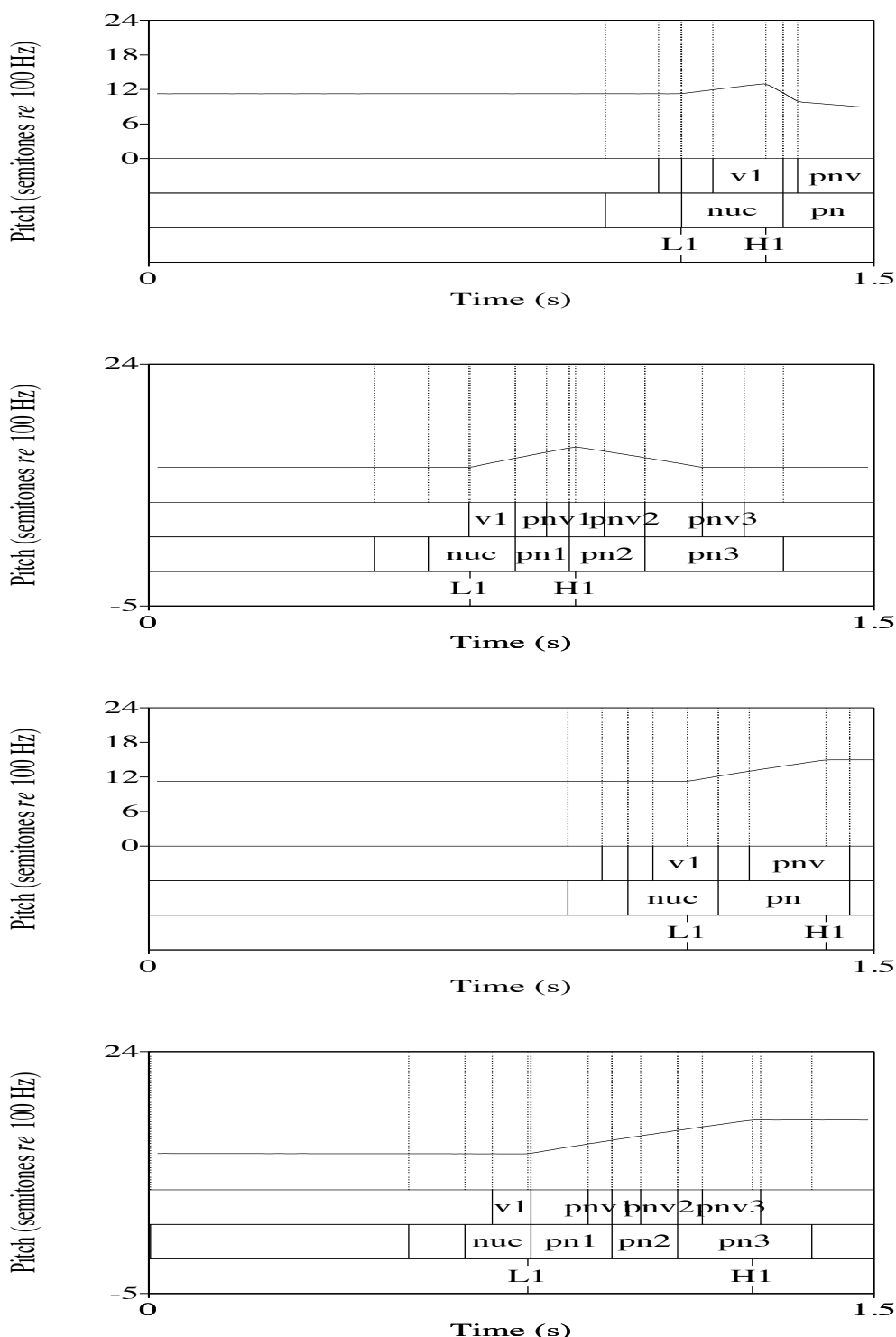
Figure 4.11: Stylised diagrams of the Glasgow and Belfast statement contours. These stylisations were made from average durations, alignment measurements and semitone values. I calculated the average syllable and segment durations separately for the Glasgow 'Short tails' (top diagram), Glasgow 'Long tails' (2nd from top), Belfast 'Short tails' (3rd from top) and Belfast 'Long tails' (bottom diagram). This is why I have not combined all four stylisations into a single diagram.

In the 'Long tails', H1 was aligned beyond Endv1 and therefore beyond the nuclear syllable. The mean proportional alignment measurement was 0.235 for the 'Long tails' statements, showing that in the 'Long tails', the distance between Endv1 and H1 takes up almost 25% of the duration of the 'Tail'. Of the Glasgow 'Long tails' with H1 beyond the nuclear syllable, 18 had H1 aligned in the first postnuclear syllable, 28 in the second postnuclear syllable and 2 in the third postnuclear syllable. Of the statements alone, 4 had H1 aligned in the first postnuclear syllable, 9 in the second postnuclear syllable and 1 in the third postnuclear syllable. It is plausible that differences in the segmental content of the target nuclear compounds may account for the alignment differences here, but an investigation of this is beyond the scope of the current thesis (c.f. Mayo et al., 1997). Microprosodic effects that were not entirely removed by smoothing of the $f0$ contour may well also account for some of the variability here (see section 4.8.7 above). The important point though is that there is a trend for H1 to be aligned even further than the first postnuclear syllable.

The difference in alignment of H1 between 'Short' and 'Long tails' (considering all sentence types together, not just statements) was significant for the Glasgow data ($pmcmc < 0.05$). What this demonstrates is synchronic allophony conditioned by prosodic context (number of following unstressed syllables). The synchronic pattern provides an explanation for how the nuclear rising phenomenon in the Glasgow 'Long tails' statements from these few speakers has developed. It gives some evidence that H1 may drift to the right with the increase in the number of following unstressed syllables in the 'Long tails'. There was also consistently a final low at the bottom of the speaker's range in the Glasgow statements and wh questions in particular. The important point about Glasgow statements[20] may be the obligatory nature of this final low. In the 'Short tails', this is realised on the single unstressed syllable following the nuclear syllable. This may result in the preceding H1 being earlier aligned to avoid tonal crowding with the final low. Overall, I characterise the Glasgow statement nuclear tone in this small dataset as a having a 'rise-fall' contour shape. The Glasgow statement 'rise-fall' bears some resemblance to the East European Question Tune (EEQT) (Grice et al., 2000; Ladd, 2008) but is not identical to the EEQT in any of the varieties in which it occurs. A feature like [+ delayed peak] (Ladd, 1983) could demonstrate that

---

[20]The more general historical hypothesis underlying the Alignment hypothesis expects "Realignment" specifically as an explanation of the presence of nuclear rises on statements as opposed to other sentence types (section 4.7.2 above).

even in the 'Long tails' set, there is a nuclear fall with a rise element at the beginning (see also Cruttenden, 1997, p. 135). As previously mentioned, the reader should caution against the assumption that the pattern reported for the few Glasgow speakers in this study must necessarily represent Glasgow English speech more generally.

Importantly, in the Belfast data, there was no difference in the alignment of H1 between the 'Short' and 'Long tails'. This was because H1 in these Belfast data was almost always aligned in the final syllable of the IP. Therefore, these Belfast data do not support the Alignment hypothesis. I assume that H1 in the Belfast data shows Segmental Anchoring (see chapter 3) to some point within the IP-final syllable in both 'Tail' types and has not 're-aligned' from any earlier point. Further confirmation of this comes from measuring the alignment of H1 from the onset of the IP-final vowel (expressed as a proportion of the duration of the final vowel (see appendix D)). There was no difference again in the alignment of H1 between 'Short' and 'Long tails'. The reason for choosing this new segmental landmark was that Schepman et al. (2006) recommend choosing nearby landmarks from which to measure alignment. H1 in Belfast is aligned much closer to the onset of the IP-final vowel than to Endv1.

4.12 Transfer hypothesis discussion

To evaluate this hypothesis, I focus on within-variety comparisons of statements, lists and questions using the parameters of alignment and scaling. The underlying question is whether Belfast and Glasgow English speakers make systematic distinctions between statement, list and question intonation (see Ladd, 2008, p. 127). Major differences between statements and lists or between statements and questions would make it difficult to support an account of speakers simply transferring typical list or question intonation onto a statement context.[21]

4.12.1 Alignment: Glasgow

There were alignment differences between statements, lists and questions in both varieties.

---

[21]I do not discuss the possibility of transfer of continuation or questioning meaning/attitude onto a statement context in this section.
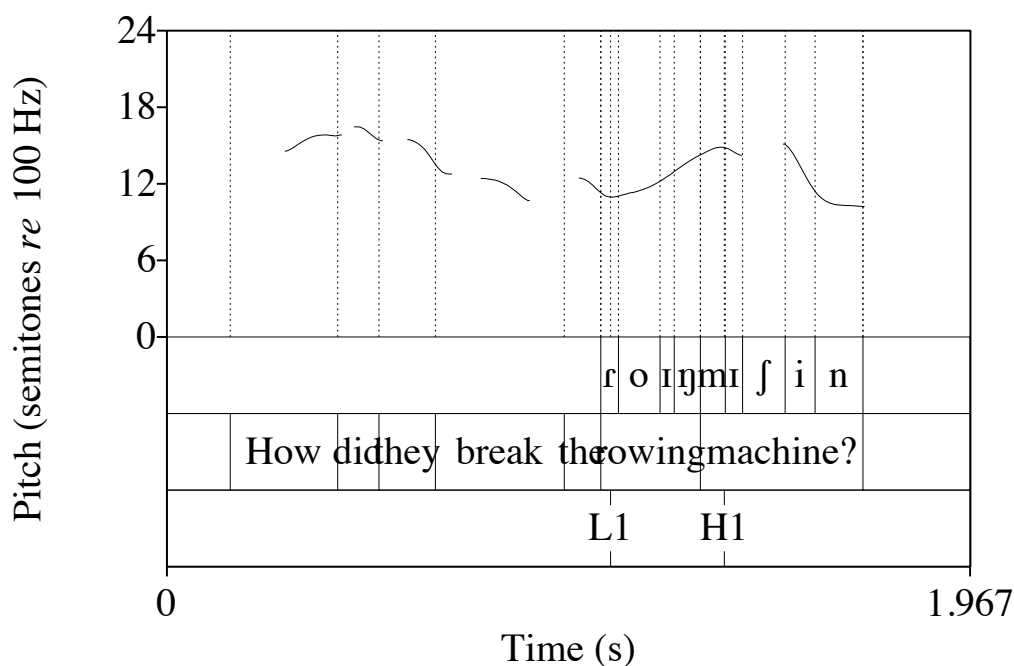
Figure 4.12: *f*0 trace of a Glasgow wh question from speaker AG. Notice its similarity, particularly in the alignment of H1, to the Glasgow statement in figure 4.8 above.

The alignment of L1 in the Glasgow data was significantly different between statements and lists (*pmcmc* < 0.001). List intonation in the Glasgow data was often produced with an extremely flat nuclear contour with a small rise at the IP boundary. L1 in the Glasgow lists was aligned much later than in the Glasgow statements (section 4.9.2 above). However, the number of usable tokens of Glasgow lists was small. The alignment of L1 in the Glasgow data tended to be earlier (with respect to Begv1) in wh questions than in statements, which in turn tended to have earlier L1 alignment than y/n questions. There was no significant difference between statements and either question type though.

The alignment of H1 in Glasgow was also significantly different between statements and lists (*pmcmc* < 0.05), with lists again showing later alignment than statements. In the alignment of H1 in Glasgow, statements showed a trend for earlier alignment (with respect to Endv1) than both question types. Wh questions had earlier alignment than y/n questions. Again, these differences between statements and questions were not significant. Wh questions, like statements, were different to lists (*pmcmc* < 0.05). The alignment of L1 and H1 between Glasgow statements and wh questions is particularly close.

4.12.2 Alignment: Belfast

The alignment of L1 in the Belfast data was significantly different from both lists ($pmcmc < 0.05$) and wh questions ($pmcmc < 0.05$). L1 was aligned earlier in wh questions than in statements, which in turn had earlier alignment than lists.

In the alignment of H1 in Belfast, statements were significantly different to both wh questions ($pmcmc < 0.05$) and y/n questions ($pmcmc < 0.001$).[22] Both question types had a consistently earlier alignment of H1 than statements. There was no difference in the alignment of H1 between the two question types. The earlier alignment of H1 in questions than statements was often accompanied by some small 'slumps' after H1 in questions. In contrast to the alignment of L1, there was no difference between statements and lists in the alignment of H1. Some previous research argues that later alignment of peaks can signal questionhood (Haan, 2002; House, 2003; Gussenhoven, 2004; Segerup and Nolan, 2006). This is clearly not supported by these Belfast data, in which questions show earlier alignment than statements. However, the scaling results showed Belfast questions behaving very much in line with previous research on intonational differences between statements and questions (section 4.12.3 below).

The alignment differences between Belfast statements and questions show that some alignment modifications would certainly be necessary to convert a Belfast question into a Belfast statement. Further, Belfast statements and lists are generally closer in alignment than Belfast statements and questions.

Overall, however, the H1 alignment differences between Belfast statements and questions are smaller than the striking difference in the alignment of H1 between Belfast and Glasgow. For example, the mean alignment of H1 in the Belfast 'Long tails' took up 84% of the duration of the 'tail' (section 4.9.3 above). In the Belfast 'Long tails' y/n questions it took up 72% of the 'tail' but in the Glasgow 'Long tails' statements it took up just 24% of the 'tail'. H1 in Belfast statements and questions is clearly associated to the IP-boundary, whereas the Glasgow H1 is probably associated to the accented syllable.

---

[22]This was the case for alignment measured from Endv1 and from the onset of the IP-final vowel.
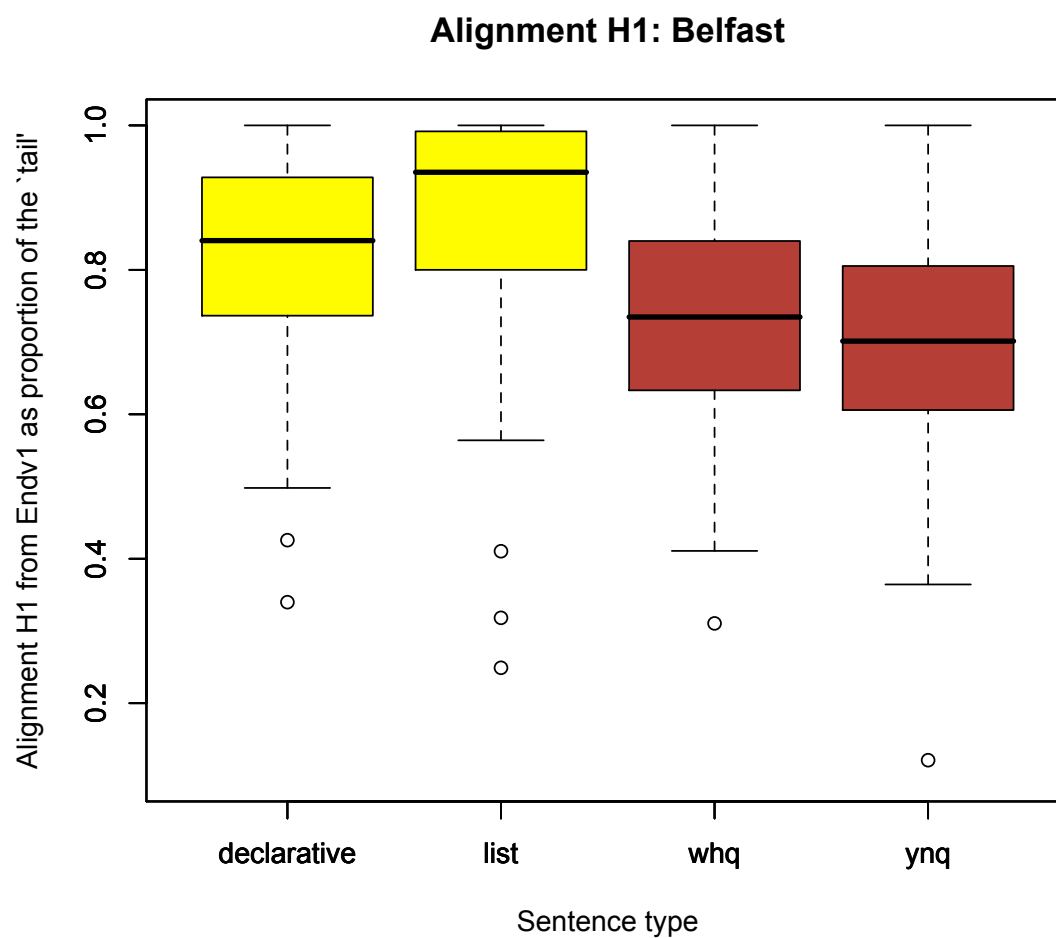
Figure 4.13: Boxplot displaying the alignment of H1 in the Belfast data. 'Declarative' refers to 'statement'. Notice that the alignment of H1, measured from Endv1, is proportionally earlier in the 'tail' in both sets of Belfast questions than in the declaratives/statements and lists.

4.12.3 L1 scaling

Now I turn to an evaluation of differences between statements and questions in relation to the parameter of scaling. Questions typically have a higher scaling of peaks and/or a wider $f0$ excursion than corresponding statements (Hayes and Lahiri, 1991; Hirst and di Cristo, 1998; Haan, 2002; Yuan et al., 2002; Gussenhoven, 2004; Makarova, 2007; Ladd, 2008).

There were no systematic differences between sentence types in either variety in the scaling of L1. Therefore, L1 in statements, lists and questions is pitched at roughly the same place with respect to the speaker's mean $f0$.

4.12.4 H1 scaling

There were strong differences though between statements and questions in relation to the scaling of H1. In both varieties, questions had a consistently higher scaling of H1 than statements, reflecting the common cross-linguistic trend. Statements were significantly different in the scaling of H1 to wh questions and y/n questions (Glasgow: statements vs. wh questions and statements vs. y/n questions, both $pmcmc < 0.05$; Belfast statements vs. wh questions and statements vs. y/n questions, both $pmcmc < 0.001$). For example, the mean normalised scaling of H1 on statements in the Belfast 'Long tails' was 1.808 semitones above the speaker's mean $f0$ (section 4.9.7 above). In wh questions, it was 4.551 semitones. In the Glasgow data, the corresponding values were 1.242 semitones for statements and 2.188 semitones for wh questions. There were significant interactions between Variety and Sentence type in relation to Statements and wh questions ($pmcmc < 0.01$) and in relation to Statements and y/n questions ($pmcmc < 0.01$). This reflects that there was a much greater difference in H1 scaling between Belfast statements and questions than between Glasgow statements and questions. In neither variety was there a difference between the two question types. This indicates a broad statement vs. question distinction without refinements of scaling between the two question types. There was also no difference in either variety between statements and lists. This stands in contrast to the clear difference between statements and both question types. Wh questions were also different in the scaling of H1 to lists (Glasgow $pmcmc < 0.01$, Belfast $pmcmc < 0.001$) as well as to statements. This reveals even more clearly that in both varieties, statements and lists are patterning separately from questions in relation to H1 scaling.
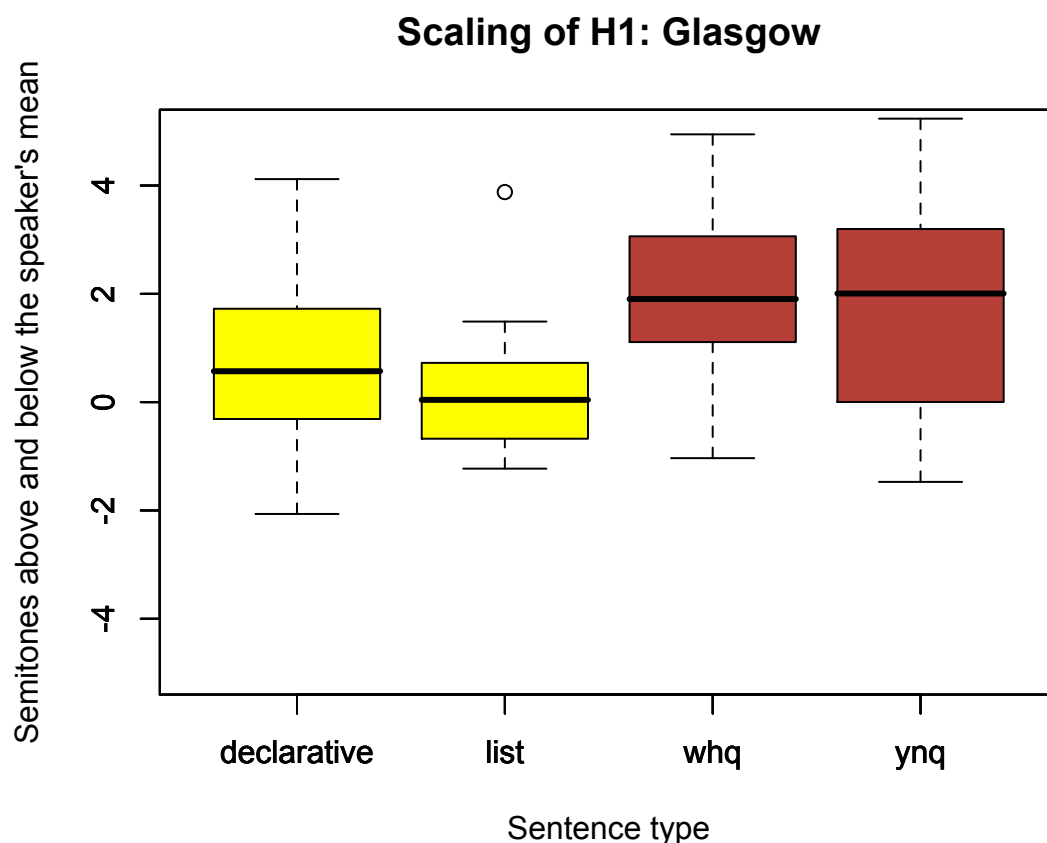
Figure 4.14: Boxplot displaying the normalised scaling of H1 in the Glasgow data. Notice the difference in H1 scaling between declaratives/statements and questions. Notice also that declaratives/statements are much closer in H1 scaling to lists than to questions.

The H1 scaling findings strongly echo findings for the $f0$ excursion (in linear Hertz) between L1 and H1, which I calculated separately. Full details are given in appendix D. In Experiment I, Belfast questions in IViE data set had a significantly different $f0$ excursion to the Belfast statements (Cinderella passage) but not to the Belfast statements (Sentences list) (section 4.4.2 above).

The H1 scaling findings put more distance between statements and questions than between statements and lists in both varieties. However, in the Glasgow data, there were strong alignment differences between statements and lists. This suggests that the Transfer hypothesis is not a satisfactory account of the Glasgow statement nuclear tone.

The Belfast statement data show some strong connections with Belfast lists in H1 alignment and scaling and thus the Transfer hypothesis looks plausible here.
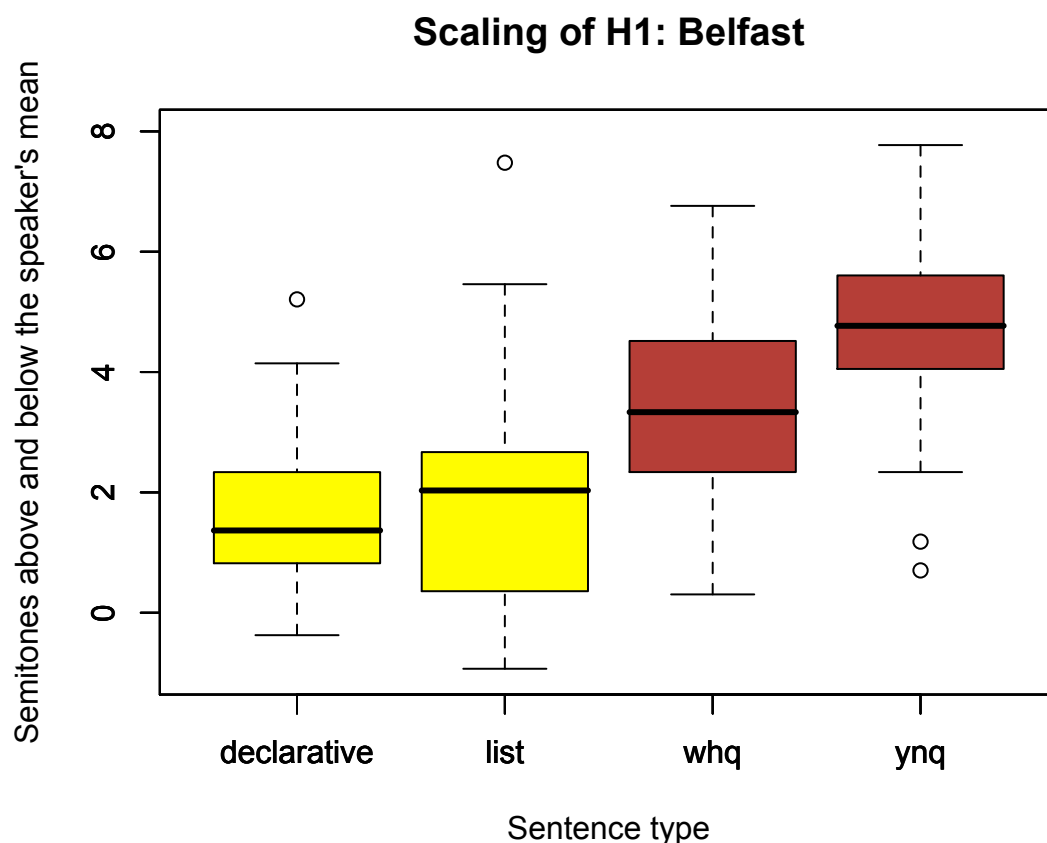
Figure 4.15: Boxplot displaying the normalised scaling of H1 in the Belfast data. As with figure 4.14 above, notice the difference between declaratives/statements and questions. It is also important that declaratives/statements are closer to lists than to questions.

Specifically, I propose that the Transfer is from list intonation and not from question intonation. Differences in pitch peak height are known to be perceptually salient (Gussenhoven, 2004; Makarova, 2007) (as are differences in pitch range (Gussenhoven and Rietveld, 1991, p.445)). I would need to account for a reduction in scaling of H1 to form a statement from a question contour in both varieties. Admittedly, the H1 scaling differences between statements and questions may be representative of paralinguistic differences in pitch range. Nevertheless, Belfast statements and lists were also closer in H1 alignment than Belfast statements and questions. I acknowledge that it would be incorrect to argue directly that the lack of significant differences between statements and lists in the scaling and alignment of H1 give support to the hypothesis that that Belfast nuclear statement intonation involves transfer from list intonation rather than from question intonation. These findings do allow me to argue that there is more phonetic

distance in H1 scaling and H1 alignment between statements and questions than between statements and lists. This makes the transfer of list intonation more plausible than the transfer of question intonation.

The phonological status of the scaling differences between statements and questions is uncertain. On the one hand, the strongly significant differences between statements and questions might suggest a categorical difference in H1 scaling. On the other hand, the H1 scaling differences between statements and questions may be gradient paralinguistic differences, with most statements clustering at one end of the scale and most questions clustering at the other end. The plots (figures 4.14 and 4.15) of H1 scaling between statements, lists and questions in both varieties show overlap between statements and questions despite the strongly significant differences between them. It is not feasible to resolve this ambiguity in this thesis. To do so would require a categorical perception test, a standard procedure for testing putative phonological contrasts. Further, applications of this procedure in an intonational context have been problematic (Ladd and Morton, 1997; Remijsen and van Heuven, 1999). This does not take away from my finding that Belfast statements and questions are more phonetically distant than Belfast statements and lists. It is phonetic distance that I use to make inferences about the source of Belfast statement nuclear intonation.

4.12.5 'Rise-plateau' vs. continued rise

Thus far I have concentrated on the implications of the results of the main parameters of alignment and scaling on the Transfer hypothesis. Now I turn to a more general consideration of overall contour shape in Belfast and Glasgow statement and question nuclear intonation, which is directly relevant to earlier work on these varieties.

Existing reports on UNB rises (which include Belfast and Glasgow rises) have distinguished their overall contour shape from the shape of rises in other varieties of English (Jarman and Cruttenden, 1976; Cruttenden, 1997; Ladd, 2008). Cruttenden (1997, p.133ff.) described the Glasgow contour as typically a rise and the Belfast contour, typically a 'rise-plateau'.[23] Ladd (2008) presents the Glasgow contour as a 'rise-plateau-slump'. The 'rise-plateau(-slump)' would contain a rise on the nuclear syllable followed by a levelling off of pitch ('plateau') and possibly a final decline in pitch ('slump'). By contrast, Cruttenden and Ladd

---

[23]See also McElholm (1986, p.32) on Derry English.
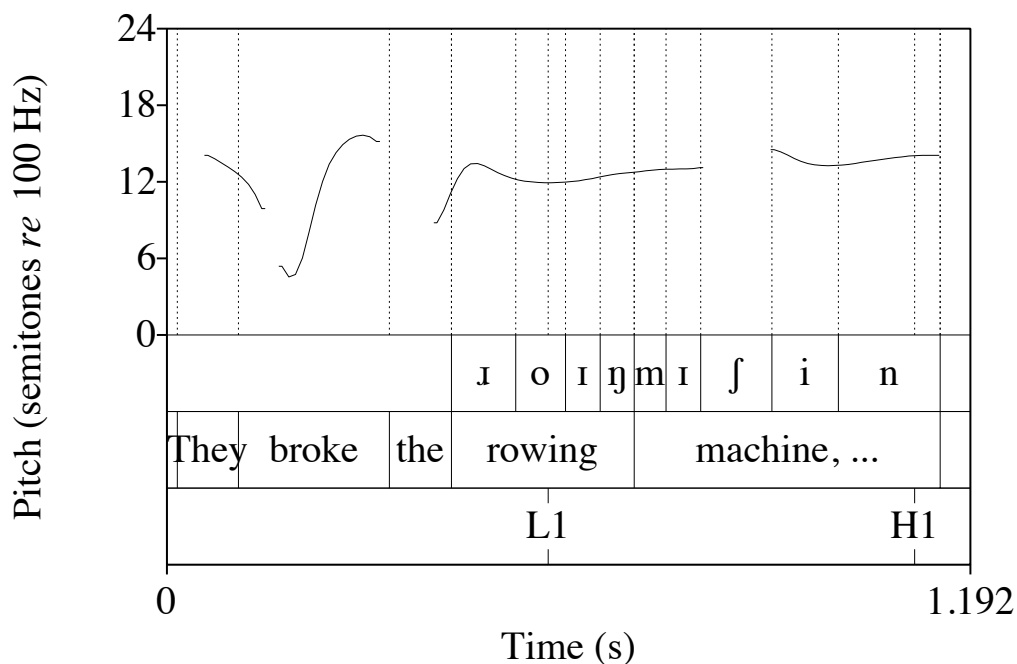
Figure 4.16: *f*0 trace of a Belfast list from speaker PT. The full utterance is 'They broke the rowing machine, robbed the jewellery store and ran to the railway station'. Notice the alignment of H1 near the end of the IP and the similar scaling of H1 to the Belfast statement in figure 4.9 above.

describe HRTs as rising continually without levelling off or declining in pitch. I have already shown that the overall nuclear contour shape of the Glasgow statement data is probably best described as a 'rise-fall' rather than a rise or a 'rise-plateau-slump' (section 4.11.2 above). Grabe (2002) makes a distinction between statements and questions in the IViE Sentences list from the Belfast English speakers. The difference between 'rise-plateau' and a continued rise in the IViE labelling system is reflected in the final boundary tone: 0% for 'rise-plateau' and H% for a continued rise. Grabe classified none of the statements (declaratives) as containing a final H% but 5.6% and 16.7% respectively of the y/n questions and declarative questions as having H%.[24]

---

[24]Gussenhoven (2004, p.301) draws on a tiny portion of the IViE data with an example of a 'rise-plateau' from Liverpool English. He claims that there would be a following fall/'slump' if there were a reporting clause afterwards. The Cinderella passage from which this is taken has several sentences with reporting clauses, so there would have been ample scope for Gussenhoven to examine these. In any case, the potential for a prosodic break between a main clause and a reporting clause means that this would not be the most appropriate example of a 'rise-plateau-slump'.
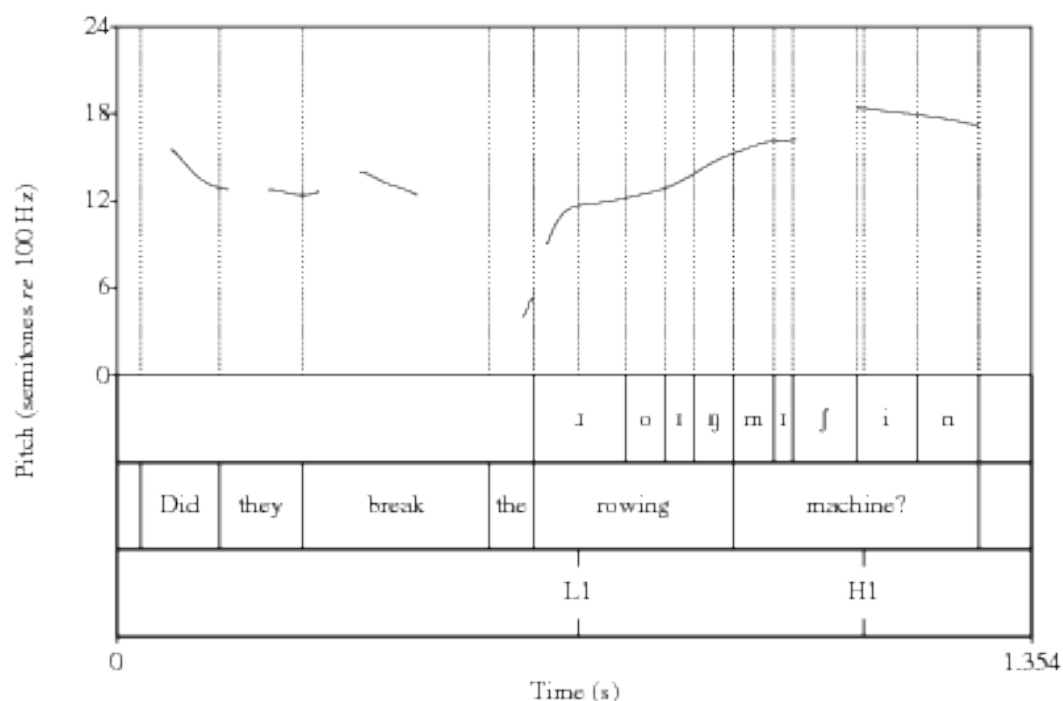
Figure 4.17: $f0$ trace of a Belfast y/n question from speaker PT. Please ignore the error around L1. Notice the slightly earlier alignment of H1 here than in the statement (figure 4.9) and list (figure 4.16) above. Also notice that H1 is scaled higher in this y/n question than in the statement or the list.

From my analysis of the IViE Belfast sentences data in Experiment I, I argue that this distinction between 'rise-plateau' and continued rise in fact needs much closer scrutiny. The IViE sentences contained words in nuclear position that were disyllabic with initial stress. One unstressed syllable after the nuclear syllable is simply not enough to evaluate the difference between a 'rise-plateau' and a continued rise. Although Lowry (2002*b*) used data from three other speech styles in the IViE corpus of Belfast English[25] and also found rise-plateaux dominating, these data still do not provide ample tokens of nuclear syllables with several following unstressed syllables. This shortcoming of the IViE data was a primary motivation for the design of the 'Long tail' test materials for Experiment II (3 or 4 following unstressed syllables).

I argue that the Belfast data from the Sentence reading task in Experiment II do not provide strong evidence of the 'rise-plateau(-slump)' contour in the way that it has previously been described. That is, the sharp rise does not just occur on the

[25]the Cinderella reading passage, the re-telling of the Cinderella passage and the Spontaneous speech data

nuclear syllable followed by a flat or falling contour over subsequent syllables. The alignment of H1 (*f*0 maximum) was located overwhelmingly in the IP-final syllable. Just one of EL's 'Long tails' had H1 aligned in the penultimate syllable, 7 of speaker PT's and none of speaker RB's. In these exceptional cases, there was a noticeable pitch slump in the final syllable and sometimes a small plateau before the slump.

Usually the pitch rose steadily from the preceding L1 across all intervening unstressed syllables without an obvious 'plateau' effect. It would not be straightforward to assess objectively whether there is another high turning point marking the beginning of some kind of 'plateau' before the true *f*0 maximum. Earlier in this chapter, I demonstrated problems with the 'elbow' scripts of Welby & Beckman (D'Imperio, 2000; Welby, 2003). This is why I did not use these scripts in the analysis of the data from Experiment II.

After H1 was reached in the final syllable in the Belfast data, there was often a very small *f*0 slump. Where this small slump was present, it was actually far more common in questions than in statements (see also Lowry, 1997). It would be extremely difficult to find a completely objective procedure for determining the presence or absence of a final small slump that is perceptible. I base my assessment here on auditory analysis combined with inspection of the *f*0 trace. In the 'Long tails', EL had slight audible final slumps in 3 statements, 7 wh questions (1 of these was the more major slump) and 4 y/n questions. PT had one slump on a statement, one on a list, 6 on wh questions and 7 on y/n questions. RB had 1 slight slump on a wh question and 2 on y/n questions. Perhaps the presence of this small slump could be related to the use of an expanded pitch range and a higher scaling of H1 in questions as opposed to statements. If H1 in questions is at the top of the speaker's pitch range, perhaps she cannot sustain this and the pitch drops slightly. I emphasise that these assessments are tentative but I argue that the Belfast statement rises in these data do not fit the expected contour shape of a 'rise-plateau(-slump)'. If a 'slump' occurs, it is more likely in fact to occur in questions rather than statements, though the question nuclear contour shape in such instances would be more appropriately termed 'rise-slump' than 'rise-plateau-slump'. The finding of earlier alignment of H1 in questions than statements also supports the position that there is more room for a following 'slump' in questions than statements. Apart from the trend of the 'slump' occurring in questions rather than statements, it was difficult to find any

further consistent pattern. Perhaps it is optional because it is such a small element occurring at the very end of the IP (see also Lowry, 1997). Though Ladd (2008, p.129 note 18) assumes the presence of a 'plateau', he still raises the crucial point that the difference between final level pitch (perhaps we could substitute 'slump' instead of 'level') and final rising pitch may not be as clear in UNB varieties as in other varieties of English.

I acknowledge the possibility that these Belfast English speakers may have modified their nuclear rises in the few years that they have not lived in Belfast. Perhaps under exposure to more standard varieties of English, including the phenomenon of High Rising Terminals (HRTs), the 'rise-plateau' has become realised as a continued rise (a HRT) in their speech. Three speakers is of course a very small sample size too.

Another relevant argument is that the use of alignment measurements from just two points (L1 and H1) was too sparse to capture the potentially more complex shape of the Belfast English rise. Perhaps there is some kind of plateau which begins before the $f0$ maximum but it is impossible to capture this using only maxima and minima as measurements.

There are also other potential reasons for the absence of a clear plateau shape in these Belfast English data. The first of these is the presence of secondary stress on the final syllable of some of the 'Long tails' target words/compounds. This could be a confound in four of the ten 'Long tails' target words/compounds used in the analysis: 'vending machine', 'rowing machine', 'volleyball court' and 'amusement arcade'. The second reason is the presence of a high vowel in the final syllable. High vowels have intrinsically higher $f0$ than other vowels, other issues being equal. This could also be a confound in four of the ten 'long tails' target words/compounds: 'vending machine', 'rowing machine', 'war memorial'[26] and 'bungee jumping'. The third reason relates to the presence of voiceless obstruents around the final syllable. These lead to an elevated $f0$ once voicing begins again. This could have been a problem in eight of the ten target words/compounds: 'vending machine', 'rowing machine', 'animator', 'volleyball court', 'dairy cattle', 'gladiator', 'bungee jumping' and 'amusement arcade'. The overall point is that these three issues may bias the location of the $f0$ maximum towards the final syllable of the IP.

---

[26]As previously explained, I essentially consider the high vowel /i/ as part of the final syllable in this compound because the final vowel was usually very reduced.

The previous description of the Belfast contour as a 'rise-plateau' was based almost entirely on auditory analysis. Using acoustic analysis, this thesis reveals a mismatch between that previous description and the rises without clear plateaux that dominated the current data set. I have just outlined several potential reasons for this mismatch. Further work is needed to decide between two main possibilities:

- The traditional description of the Belfast contour as a 'rise-plateau' needs to be revised, as the current results suggest;
- The 'rise-plateau' does indeed have robust acoustic correlates, but the nature of the speakers, test materials and measurements of the present study prevented these from being discovered.

A relevant but unanswerable question for the present is whether question rises from Standard English varieties and/or HRTs might also have small slumps in the final syllable.[27] It would also be very valuable to compare these Belfast data with Donegal Irish and Donegal English data (Dalton and Ní Chasaide, 2005; Dalton, 2007; Kalaldeh et al., 2009), as researchers of these varieties claim that 'rise-plateaux' are in evidence there.

Returning to the comparison of Belfast and Glasgow English, I emphasise that the Belfast 'slump' is not the same as the final low in the Glasgow data. The 'slump' represents a slight decline in pitch but the fall element in the rise-falls in these Glasgow data reached the bottom of the speaker's pitch range.

4.12.6   Edinburgh English wh questions

The analysis of the Glasgow English data has shown a very similar alignment of H1 between statements and wh questions (section 4.12.1 above). Impressionistically, the alignment of H1 in the Edinburgh English wh questions also appeared to be close to the Glasgow data though possibly earlier aligned. This led me to include the Edinburgh English data in a formal comparison of wh question realisation between Edinburgh and Glasgow. The mean alignment values and

---

[27]Warren (2005) and Warren and Daly (2005) found that questions had earlier alignment of the beginning of the rise than statements. In New Zealand English HRTs, they suggest a change in progress to distinguish questions from statements. The Belfast speakers in Experiment II distinguish statements from questions through H1 scaling, something which may not be done by New Zealand English speakers. Warren (2005) also reports tentatively on the perception of HRTs in New Zealand English. A small-scale study showed that people were more likely to perceive a HRT as a question when H was realised as a plateau than when it kept rising until the end of the IP. Warren concludes that the shape of the rise is perceptually important.

Alignment H1: Edinburgh and Glasgow wh questions
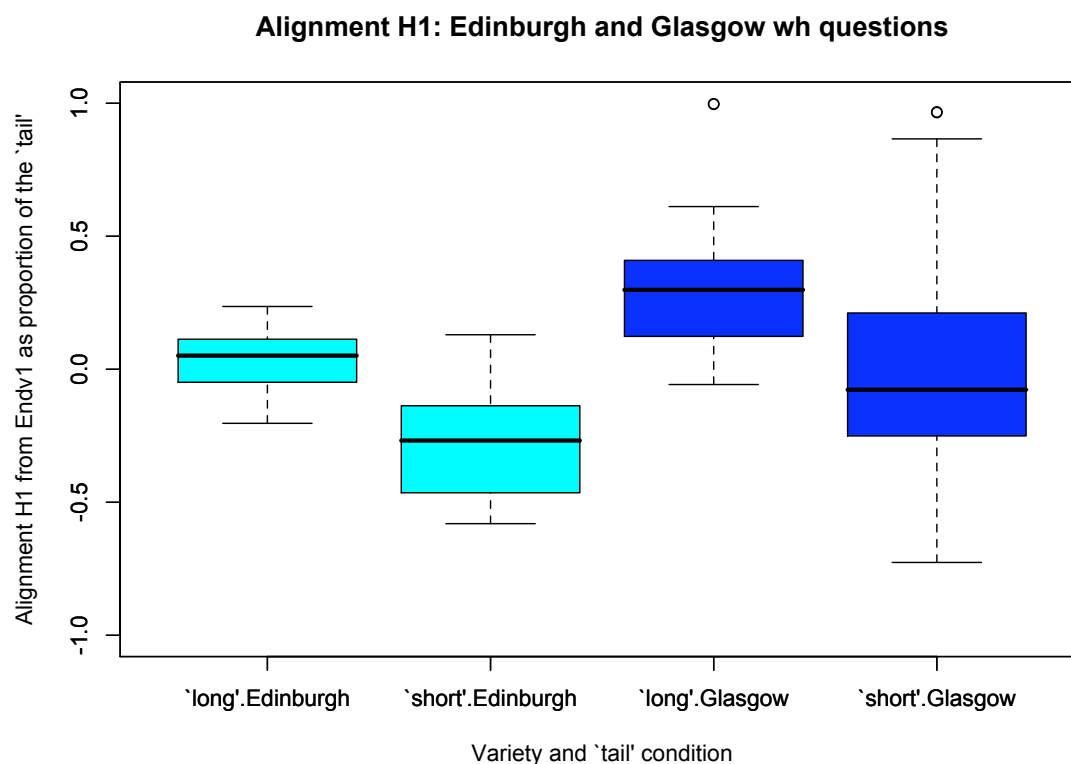
Figure 4.18: Boxplot displaying the alignment of H1 in the 'Short' and 'Long tails' in the Edinburgh and Glasgow wh questions. Notice the difference between the two 'tail' types in both varieties and that the alignment of H1 is generally proportionally earlier in the Edinburgh data. Minus values indicate that the alignment of H1 occurred before Endv1.

other statistical details are in appendix D. The Glasgow data did indeed show later proportional alignment of H1 than the Edinburgh data. The difference between the varieties was significant ($pmcmc < 0.05$). There was also an effect of 'Tail' ($pmcmc < 0.01$), with later alignment of H1 in the 'Long tails'. Unlike the comparison of Glasgow and Belfast in section 4.11.4 above, there was no significant interaction between Variety and 'Tail' here. This shows that although the Edinburgh English wh questions have earlier alignment of H1 than Glasgow wh questions, the Edinburgh wh questions behave very similarly to the Glasgow wh questions (and statements) in the 'Short' and 'Long tails' (section 4.11.4 above). Both display allophony conditioned by the number of following unstressed syllables. There were no differences between the Glasgow and Edinburgh wh questions or between the 'Short' and 'Long tails' sets along any of the other parameters, however.

4.12.7   Interim conclusion

To summarise, the analysis of the data from the Sentence reading task leads me to argue that the Glasgow data support the Alignment hypothesis. The evidence for this comes from the allophony in the 'Tail' condition. There was later alignment of H1 in the 'Long tails' than the 'Short tails' (section 4.11.4 above). The Glasgow wh questions are very close to the Glasgow statements in alignment and also display the alignment difference between 'Short' and 'Long tails'. Thus Glasgow wh question intonation may also have come about through "Re-alignment" of H1 beyond the nuclear syllable. This may also be the case for the Edinburgh wh questions (section 4.12.6 above). The alignment connection between Glasgow statements and wh questions may also support an account of Glasgow statement intonation involving transfer of wh question intonation. Wh questions had a higher scaling of H1 than statements, however.

The Belfast data do not show support for the Alignment hypothesis because there was no difference in the alignment of H1 between 'Short' and 'Long tails' there. I argue that they support the Transfer hypothesis, with the transfer coming from lists rather than from questions. There was more distance between statements and questions than between statements and lists in relation to the alignment of H1 and also the scaling of H1. Belfast statements and lists were different in relation to the alignment of L1, but I believe this is outweighed by their connections along the other parameters.

The most important implication of this is that I propose different origins for Glasgow and Belfast nuclear statement intonation from the small set of speakers whose data were at my disposal. This goes against the tradition of considering the statement rises in these varieties to be part of the same overall phenomenon (Cruttenden, 1997). The acoustic phonetic analysis which I conducted offers this fresh perspective. Cruttenden's work was based on auditory analysis and at least at the early stages of his work (Jarman and Cruttenden, 1976), he would not have had access to the kind of software that is widely available for acoustic phonetic analysis today. This acoustic phonetic analysis of contemporary data cannot give evidence of the actual trajectory of the development of nuclear statement intonation in Belfast and Glasgow. What it does show is that synchronically, there is a clear difference in the alignment of H1 between the two varieties. Within each variety, the alignment of H1 behaves very differently in relation to the contrast between 'Short' and 'Long tails'. H1 in Belfast remains aligned in the IP final

syllable in both conditions. By contrast, H1 in Glasgow is aligned within the nuclear syllable in the 'Short tails' but in a postnuclear syllable in the 'Long tails.' Admittedly, what the Glasgow data show is a pattern of stable synchronic allophonic variation. Nevertheless, the synchronic differences between Glasgow and Belfast fit with different plausible diachronic trajectories for the small set of speakers from the two varieties respectively. They also fit with different respective phonological analyses.

The potentially different origins for Glasgow and Belfast nuclear statement intonation suggest different mechanisms of intonational change. The Alignment hypothesis is based on the notion of gradual phonetic change which may become categorical/phonologised once the peak moves beyond the nuclear syllable. The Transfer hypothesis posits that intonational change is not primarily phonetic but rather pragmatic. The idea is that Belfast English speakers may have chosen to use continuation/list intonation as the canonical nuclear intonation pattern for statements, while not necessarily preserving the continuation function. Speakers of other varieties of English have not rooted statement intonation in the category of continuation. This proposal resembles the account of Bolinger (1978) for the appearance of rising intonation on statements in UNB varieties. The major difference is that he believed that questions were the source of the transfer rather than continuation. Hognestad (2006, p.115) also believed that pragmatics could influence lexical pitch accent/intonational change in Scandinavian.

4.13   Map task

The comparisons of Belfast and Glasgow English have thus far only considered the results of the Sentence reading task. The Map task data offer unique further possibilities for assessing the two primary hypotheses.

4.13.1   *Within-variety comparisons*

I chose mainly to analyse the Belfast and Glasgow Map task data separately. There were two main reasons for this. The first is that I noticed two patterns of H1 alignment in the non-questions in the Glasgow data. One pattern was similar to that of the Sentence reading task, in which H1 was aligned in the nuclear syllable ('Short tails') or beyond it ('Long tails'). The other pattern showed H1 aligning in the IP-final syllable, rather like the typical Belfast pattern of the Sentence reading task. The second reason was that it proved more problematic for

me to get enough Glasgow Map task data without $f0$ errors at the crucial measurement points than for the Belfast data.

The primary goal of the within-variety Glasgow Map task analysis was to try to probe the two above mentioned patterns a bit more systematically. The existence of these two patterns may give another perspective on the Sentence reading task pattern as indeed being distinct from the Belfast Sentence reading task pattern.

There were two goals to the within-variety Belfast Map task analysis. The first was to examine if there was a difference in the alignment of H1 between 'Short' and 'Long tails'. This addresses the Alignment hypothesis (section 4.11 above). The second was to examine if there were differences in alignment and scaling between questions and non-questions. This addresses the Transfer hypothesis, although it does not allow for a counterpart to the 'list' intonation of the Sentence reading task.

Mean values for alignment and scaling and other statistical details are in appendix D.

4.13.2 Glasgow results

Due to my interest in the two different nuclear contours that appeared to be in the Glasgow data, I added a separate fixed factor to the Glasgow Map task analysis: Contour type (final low vs. no final low). There was a difference in the alignment of H1 between the two contour types ($pmcmc < 0.05$). The alignment of H1 (measured from Endv1 as a proportion of the 'tail') was later as expected in those utterances with no final low than in those with a final low. This echoes the difference between Glasgow and Belfast in the alignment of H1 that I found in the Sentence reading task (section 4.11.1 above). There were no differences between the two contour types in any of the other parameters. There were also no differences between 'Short' and 'Long tails' or between 'non-questions' and 'questions' in any of the parameters.

I expected an interaction between 'Tail' and 'Contour type'. This was because I expected a greater difference in the alignment of H1 between the 'Short' and 'Long tails' in the utterances with a final low than in those with no final low. This would form a parallel to the interaction between Variety and 'Tail' that I found in section 4.11.1 above. It would show that the utterances with a final low were clearly behaving differently (allophony between 'Short' and 'Long tails')

Alignment H1: Glasgow Map Task
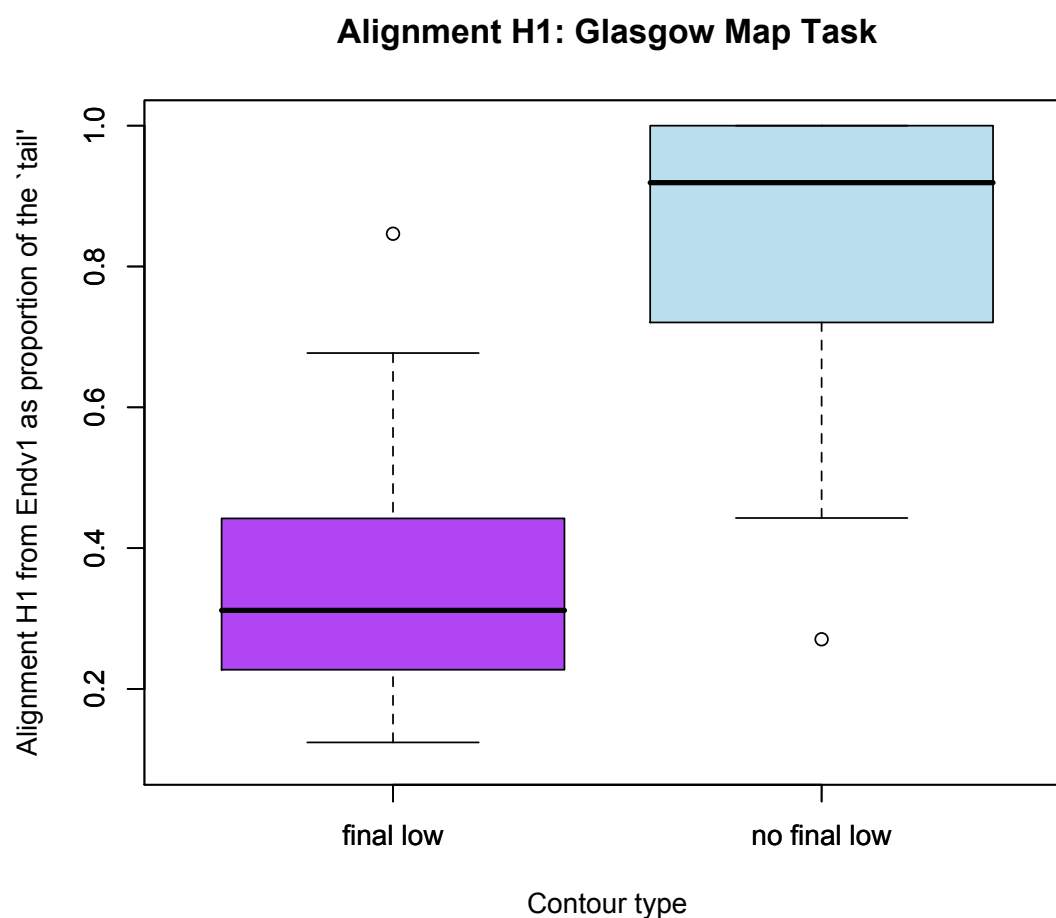
Figure 4.19: Boxplot displaying the alignment of H1 in the Glasgow Map task utterances with and without a final low. Notice that the alignment of H1, measured from Endv1, is proportionally much later in the utterances without a final low than in those with a final low.

in relation to the alignment of H1 than those with no final low (no allophony). However, there was no significant interaction. There does indeed seem to be some kind of distinction though, between a pattern like that found in the Glasgow Sentence reading task and one like that found in the Belfast Sentence reading task. If these Glasgow speakers do use two distinct nuclear contours, it does provide further support for the diachronic and phonological separateness of the Glasgow and Belfast nuclear contours from my small set of speakers that I argued in relation to the Sentence reading task (section 4.12.7 above).

The absence of any H1 scaling difference between non-questions and questions in these Glasgow Map task data goes against the typical finding of higher H1
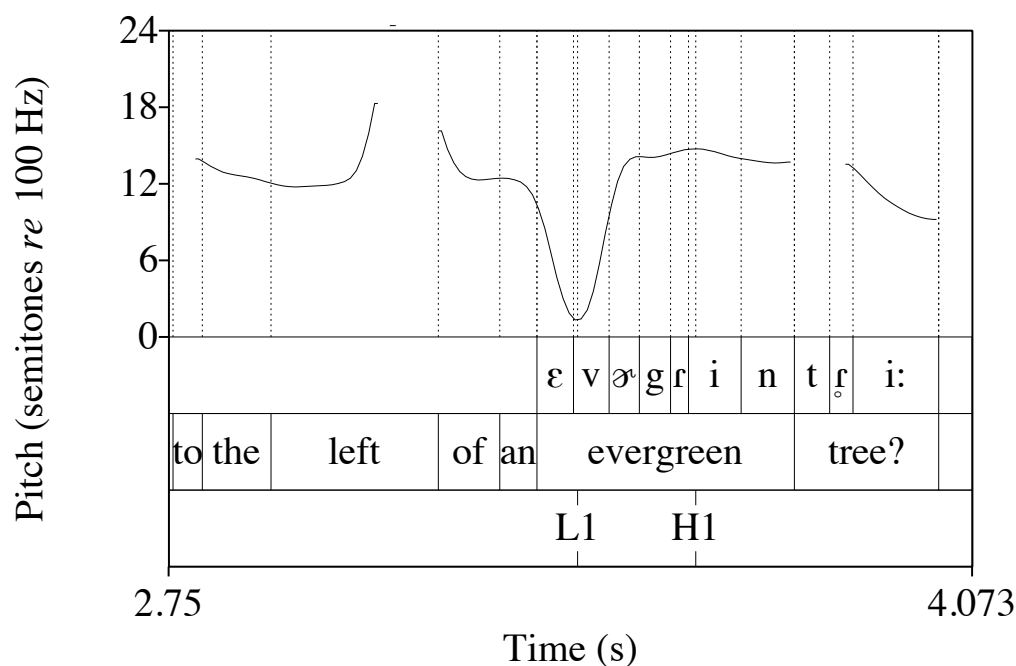
Figure 4.20: $f0$ trace of nuclear intonation with a 'final low' from speaker AG (Instruction Giver) in the Glasgow Map task. The full utterance is 'Do you have the start point marked, ehmm, above and to the left of an evergreen tree?' Please ignore the error around L1. The important points are the alignment of H1 beyond the primary stressed syllable in 'evergreen' and the noticeable fall in $f0$ afterwards.

scaling in questions, which I did find in the Glasgow Sentence reading task data (section 4.12.4 above).

The utterances without a final low in non-questions and questions often seemed to have the function of 'are you following me?' (see figure 4.21).[28] This is a common function of High Rising Terminals (HRTs/'Uptalk') noted elsewhere in the English-speaking world (Ladd, 2008, p.125). It is possible that these Glasgow utterances show that HRTs have spread to young female speakers of Glasgow English. Thus HRT nuclear intonation and typical Belfast nuclear intonation may have close connections which deserve future study. This perspective is extremely important as until now, UNB statement intonation in general has been considered quite separate from HRTs (Cruttenden, 1995, 1997; Ladd, 2008). This is not to say that I consider that the Belfast statement nuclear rise in the Sentence Reading task is a necessarily a HRT. As the data in the Sentence reading

---

[28]It is not within the scope of this thesis to discuss discourse function and intonation on any large scale. This is why I am tentative on this point.
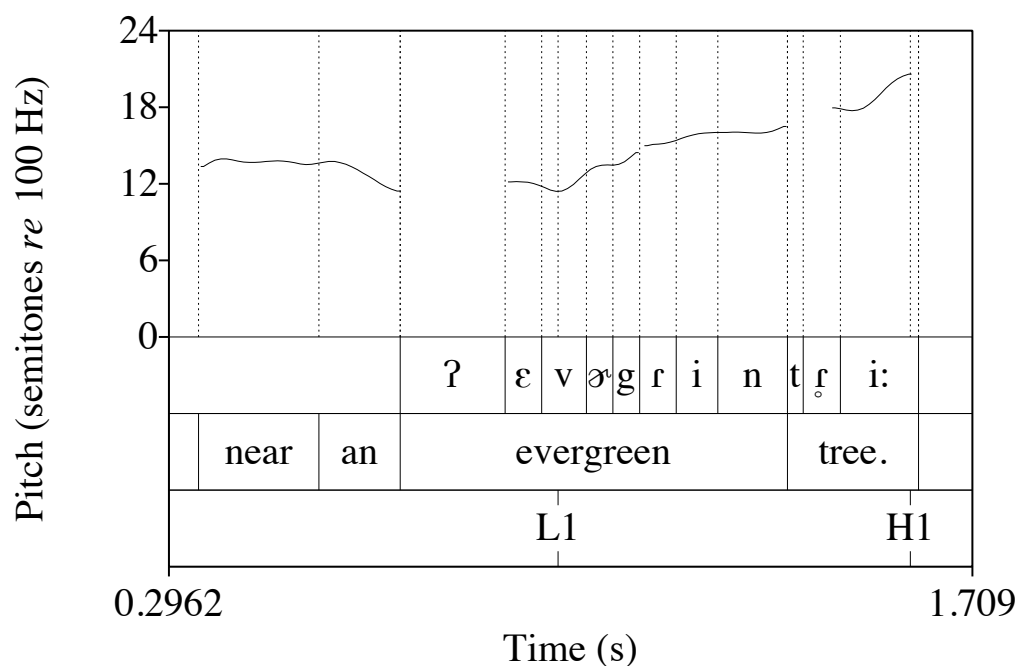
Figure 4.21: *f*0 trace of nuclear intonation without a 'final low' from speaker AG (Instruction Giver) in the Glasgow Map task. The full utterance is 'So the starting point is, ehmm, near an evergreen tree.' Contrast the alignment of H1 in the final syllable in this utterance with its alignment in the previous figure.

task show, the Belfast nuclear contours do not just occur with the 'are you following me' function. There may be a difference in the alignment of L1 between the Belfast/Glasgow 'HRT' rises and HRTs elsewhere. For example, female New Zealand English speakers begin their rises in a poststress syllable rather than on the stressed syllable itself (Fletcher et al., 2005).

4.13.3   *Belfast results*

There was no difference in the alignment of H1 between 'Short' and 'Long tails' in the Belfast data. This forms a parallel to the findings from the Sentence reading task (section 4.11.4 above) in which there was also no difference between the two 'Tail' types in the Belfast data. Therefore, these Belfast Map task data do not support the Alignment hypothesis. There were no differences in the Belfast Map task data between the 'Short' and 'Long tails' in any of the other parameters.

In relation to the Transfer hypothesis, there were differences in scaling of L1 and H1 between non-questions and questions. These differences were significant (L1

Scaling of H1: Belfast map task
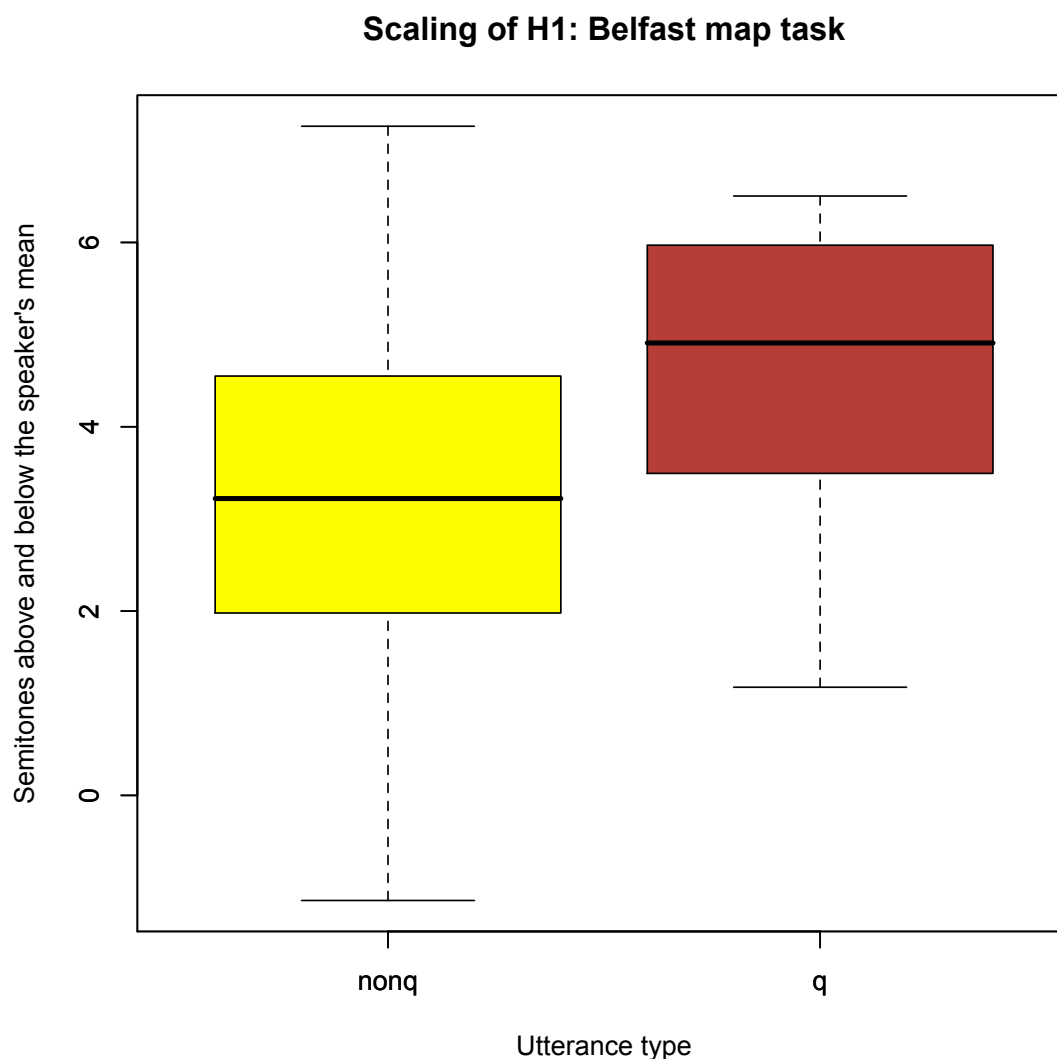
Figure 4.22: Boxplot displaying the normalised scaling of H1 in non-questions and questions in the Belfast Map task data. Notice the higher overall scaling of H1 in the questions than the non-questions.

scaling $pmcmc < 0.05$, H1 scaling $pmcmc < 0.01$). In these Belfast Map task data, L1 was higher scaled with respect to the speaker's mean pitch in the questions than in the non-questions. I did not find a difference in L1 scaling between Sentence types in the Sentence reading task. As expected, the normalised scaling of H1 was also higher in questions than non-questions. The Sentence reading task data also showed higher scaling of H1 in both question types than in statements. There were no differences between non-questions and questions in the alignment of L1 or H1 (or in 'nuclear tone rate').
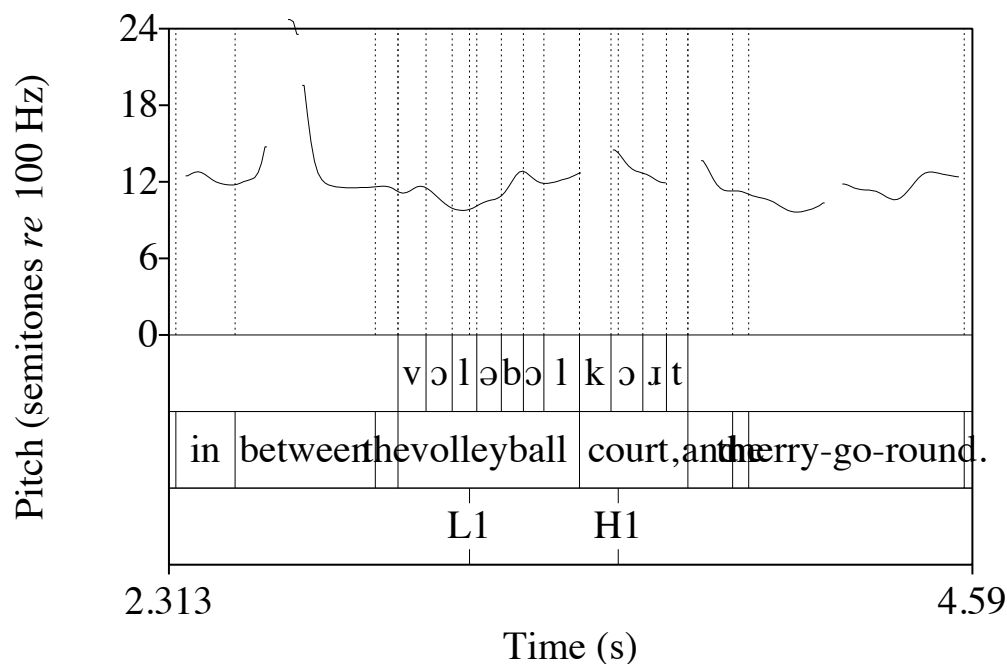
Figure 4.23: *f*0 trace of a Belfast Map task non-question from speaker EL (Instruction Giver). The full utterance is 'But then go in between, up, in between the volleyball court% and the merry go round.' There is an IP break after 'volleyball court', marked here by %.

Overall, the Belfast Map task data allow me to continue the arguments I have made from analysis of the Belfast Sentence reading task data. The Belfast Map task data do not support the Alignment hypothesis as there was no difference between the 'Short' and 'Long tails' in the alignment of H1. H1 was overwhelmingly aligned in the IP-final syllable, as in the Sentence reading task. The scaling differences between non-questions and questions may provide further support for my argument that statement intonation in Belfast has not involved transfer from question intonation. In the Sentence reading task, the Belfast statements were closer in their scaling and alignment of H1 to the lists than to questions, suggesting that transfer from list intonation is more likely. Even accepting the paralinguistic status of pitch range that may underlie the scaling differences between statements and questions, the Belfast data suggest that some kind of transfer is more plausible than an alignment change like that evident in the Glasgow Sentence reading task data.
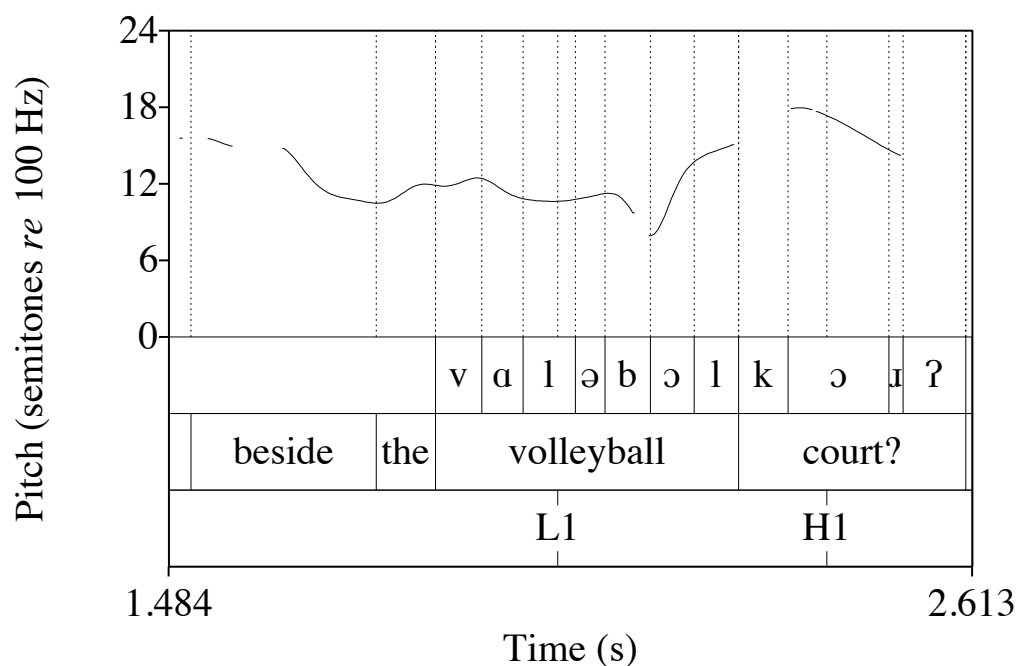
Figure 4.24: *f*0 trace of a Belfast Map task question from speaker EL (Instruction Giver). The full utterance is 'Do you have a war memorial beside the volleyball court?' Notice that H1 is higher scaled in this question than in the non-question in figure 4.23 above.

4.13.4 Between-variety comparisons

I also attempted some between-variety comparisons, comparing the Glasgow plain HRT-style rises against the Belfast rises. For none of the parameters of alignment and scaling did Variety show a significant main effect. This indicates that there are no systematic differences between these Glasgow nuclear contours without the final low and the Belfast ones along these parameters. These similarities between Glasgow and Belfast suggest that it is not necessary for me to make separate comparisons between the Belfast nuclear contours and the Glasgow contours with the final low. I assume that the results of these comparisons would be roughly the same as the within-variety comparisons by Contour shape in the Glasgow data.

One argument would be that these Glasgow nuclear contours without the final low are not in fact HRTs but actually UNB rises which had failed to emerge in the Sentence Reading task. Though I suspect that greater work on the discourse

function aspect would indeed confirm the HRT discourse functions in these Glasgow rises without the final low in contrast to the contours with the final low, it is difficult to discount this possibility. Nevertheless, it does not obscure two key points:

1. There is a difference in the Sentence reading task between these small sets of Glasgow and Belfast speakers.
2. Nuclear contours without a final low do not appear to have the same origin as those with a final low. In addition, they appear to be phonologically distinct (section 4.11.1 above).

4.14   Spectral tilt and Peak amplitude

The primary parameters I have discussed so far have been parameters to do with $f0$/pitch. I also explored a few secondary parameters related to Spectral tilt and Peak amplitude. The reason for including these secondary parameters was to examine the Transfer hypothesis further. The underlying question was whether speakers distinguished statements, lists and questions along these parameters as well as or instead of using $f0$ alignment and scaling. These secondary parameters would also have enabled comparisons between the Belfast and Glasgow data.

To measure Spectral tilt, I used scripts written by Tim Mills and discussed in Mills (2008). The measure I chose was the A1-A2 measure (see Guion et al., 2004). I measured spectral tilt on both vowels in the target nonsense words 'lala' and 'vava' and was interested if there was a difference between the initial stressed vowel and following unstressed vowel. However, the results were inconclusive. There was also the problem of glottalisation in my data. Campbell and Beckman (1997) suggest that this may make some of their own results dubious.

Separate to the spectral tilt measurements, I was also interested in overall intensity. I extracted a value of the Peak amplitude (Mills, 2008) of the nuclear vowel, postnuclear vowel and a preceding reference vowel in the 'lala' and 'vava' utterances. I wished to get a measure of the difference between the Peak amplitude of the nuclear vowel with that of the postnuclear vowel and examine if there could be differences between Sentence types in this regard.

It is crucial when taking amplitude measurements on a target vowel, that these are compared with a reference sound pressure value i.e. on another vowel (Mills, 2008). I designed the sentences with the goal of eliciting a prenuclear accent on

a specific word earlier in the sentence which contained the same low vowel as the /a/ in 'lala' and 'vava' e.g. She walked on the **path** in 'lala'. The vowel in this prenuclear word was the target vowel. Unfortunately, the word with the reference vowel was often produced without a prenuclear pitch accent in wh questions and y/n questions, making much of the data unusable.

The problems with the analysis of the spectral tilt and peak amplitude measurements mean that unfortunately I am not in a position to discuss their possible contribution to evaluating the Transfer hypothesis further or to assessing potential differences between the two varieties in this regard. However, refining these measurements offers an original avenue for future research. Spectral tilt and intensity measurements have rarely been made on data like Belfast English data, in which the $f0$ of unstressed syllables is generally higher than that of stressed syllables (exception Kochanski et al., 2005).

4.15 Approaching Chapter 5

This chapter gave details of the methodology, hypotheses and results of a pilot and main experiment (Experiments I and II) involving intonational analysis of contemporary data. The hypotheses were on potential origins of nuclear statement intonation in Belfast and Glasgow English and I evaluated them using intonational distance measurements of alignment and scaling. This chapter has addressed how far acoustic phonetic analysis of contemporary data can enable the assessment of historical hypotheses in intonation. It was primarily the quantitative difference in the alignment of H1 between Glasgow and Belfast in the 'Short' and 'Long tails' conditions, combined with the presence of a final low in the Glasgow data, that strongly indicated that Glasgow and Belfast nuclear statement intonation are different from each other, based on my small set of speakers. These differences suggest different historical trajectories as well as phonological distinctness of Glasgow and Belfast nuclear statement intonation. The analyses showed that the Glasgow data supported the Alignment hypothesis whereas the Belfast data did not. Quantification along the alignment and scaling parameters enabled me to argue for the Transfer hypothesis in the case of Belfast, from the comparisons of statements, lists and questions.

Now I need to look at the role of these phonetic measurements in an historical context more generally. Do they form a valid parallel to the phonetic distance

measurements that are increasingly used to address historical questions in segmental phonetics/phonology? Is it possible or appropriate to combine the intonational measurements of alignment and scaling into an overall measure of intonational distance? Would such a combined measure work with the Phylogenetic network techniques like NeighborNet? Finally, I ask whether new approaches to segmental phonetic distance measurements emerge in the light of what I have learned from applying intonational measurements. These issues represent the prominent themes of chapter 5.

CHAPTER 5

Segments and intonation together

5.1 Introduction

One of the primary goals of this thesis is to begin to extend phonetic distance measurements from segments to intonation. In this attempt I use an historical context for the distance measurements, as this context has been to the fore in relation to segmental measures. Chapter 2 explored the details and problems of some segmental phonetic distance measures. Chapter 3 laid out a theoretical account of potential intonational change. Chapter 4 dealt with acoustic phonetic analysis of contemporary intonational data, using standard continuous acoustic parameters as a parallel to the features used in segmental phonetic distance measures. In this chapter, I bring the segmental and intonational strands together. I discuss how much of the segmental framework is applicable to intonation and then how the new inclusion of intonation in this measurement context could aid future studies of phonetic distances in segments. There are three main sections:

1. A comparison of segmental phonetic distance measurements with the intonational measurements I have used;
2. An exploration of the problems of trying to create an overall measure of intonational distance and an assessment of whether intonational measurements could work in a phylogenetic framework;
3. Suggestions on what researchers on segmental phonetic measures could learn from intonational measurements.

The increasingly multidisciplinary nature of historical linguistic study means that distance measures in various domains of language (admittedly not intonation so far) have become of great interest to non-linguists (especially anthropologists, archaeologists and evolutionary biologists). However, there is currently deep conflict between the needs of these non-linguists and current trends in linguistic (especially phonetic) theories. I offer some comments on this important problem.

5.2 Segmental and intonational distance approaches compared

5.2.1 Context of Historical linguistics

Historical linguistics is an important context for segmental phonetic distance measurements (see chapter 2). I chose to begin intonational distance measurements also with underlying historical questions about intonational change. But is the context really equivalent between segments and intonation? Apart from obvious discrepancies between segments and intonation in relation to time depth, there is a deeper issue. Segmental phonetic measures are generally designed to confirm what we already know. Researchers have methods to identify cognates and to produce language trees that represent well-established families. They also expect to be able to distinguish familial connections from contact-induced connections. Even at the level of synchronic varieties, researchers expect their measure to capture known historical connections (McMahon et al., 2007). In chapter 2, I revealed the influence of Swadesh's technique of Lexicostatistics on segmental phonetic distance measures. Heggarty (2010) has identified that Swadesh's techniques have been used to try to answer the separate questions of how close two languages/varieties are related and whether indeed two languages are related. When delving into the unknown, researchers have used phonetic distances to ask whether certain languages are related (Baxter and Manaster-Ramer, 2000; Kessler, 2001, 2007; Kessler and Lehtonen, 2006). Yet it is still unclear whether segmental phonetic distances can shed light on this issue. More relevant, researchers have not tended to use segmental phonetic distance measures to study particular trajectories of phonetic or phonological change of individual segments.

By contrast, I used intonational measurements to examine if they could help us uncover new perspectives on the historical development of intonational statement rises in Belfast and Glasgow English. Therefore, I was not trying to use

measurements to confirm the putative historical link made between these varieties by Cruttenden (1997) but rather to probe it. Perhaps ironically, the historical context in which I have set the intonational measurements may have more in common with quantification in domains outside segmental phonetics. Dunn et al. (2005), at heart, deals with the question of whether typological features can reveal anything about language history at deeper time depths than possible with lexical or phonetic distances. A parallel question for me was whether intonational parameters could illuminate historical trajectories in intonation. McMahon et al. (2005) and Heggarty (2010) use lexical distances and the NeighborNet algorithm to ask whether Quechua and Aymara are related or whether the similarities between them are more likely to be contact-induced. McMahon et al. use the different patterns of separate sets of lexical items to conclude that contact is the more likely explanation. It was differing patterns between subsets of my data that also led me to argue the position that the Belfast and Glasgow nuclear statement contours from my small set of speakers had separate origins. The intonational work has links with this typological and lexical research through the shared absence of a well-established historical framework underpinning them. A historical context for intonational measurements is different to that for segmental phonetic distance measurements in that it is more interested in the trajectory of change than confirming existing connections between varieties. The intonational approach is similar to the segmental approach though in that it assumes that different patterns in the data reflect different histories.

5.2.2 Distinguishing ancestry, borrowing and chance

When researchers use phonetic distances to try to identify ancestral connections between languages/varieties, they must be careful to distinguish distances that are reflective of different kinds of changes. Two pronunciations of a given meaning might have a very small phonetic distance because they are indeed cognates. However, this small phonetic distance might also reflect borrowing or chance resemblance. There are a number of phonological processes of sound change that have been noted to have occurred independently in several different languages (Ringe et al., 2002, p.66-68). For example, the $*ti > si$ change happened independently in Greek and Finnish and results in these two languages appearing phonetically similar in this respect without any historical connection between them.

In intonation, I have steered clear of distinguishing between ancestry and contact/borrowing, because it simply does not make sense at the shallow time depths at which I can study. In fact, Cruttenden (1997) promotes contact as an account of the presence of nuclear statement rises in Belfast, Glasgow and other UNB varieties, even though he does not describe the realisation of the rises as exactly the same in both places.[1] A major goal of segmental phonetic distances is precisely to make this distinction between ancestry and borrowing (see chapter 2).

Especially with the small number of intonational elements and parameters I used, combined with the absence of any well-established model of intonational change, it is extremely difficult to distinguish accidental or chance distance from distance representing a shared connection between two varieties (see also Longobardi and Guardiano (2009, p.1684) on syntactic distance measurements). The phenomenon of homoplasy (independently developing similarities, not indicative of a shared historical connection (Ringe et al., 2002; Nakhleh, Warnow, Ringe and Evans, 2005; Nakhleh, Ringe and Warnow, 2005)) may be much more prevalent in intonation than in segmental phonetics. Theoretically, there is precious little to indicate under which conditions an intonational parameter is likely to indicate a connection between sentence types or varieties and when similarities could be independent developments. It is still important though to distinguish the cases in which there are precedents for the nature and range of variability and in which there are not.

In synchronic data, there are precedents for alignment differences between varieties and between sentence types and for scaling differences between sentence types only (see chapters 2 and 3 for extensive references). In intonation there is no firm basis for arguing that similar scaling of L1 and H1 between a pair of varieties is evidence of a historical connection between them. This is because there is an absence of studies to show the range of variability of scaling between varieties and thus no baseline of a chance level of scaling distance. Ladd (2008, section 5.2) discusses existing studies of scaling and pitch range. Only one of them included more than one variety in an individual study (Patterson, 2000).

This is different to a within-variety Sentence type distance in scaling. In such cases, there is the precedent that there can be Sentence type scaling differences. Therefore, there is some indication of the range and direction of variability e.g.

---

[1]That is not to say that contact between the two varieties has not had an impact.

higher scaling of H in questions (Haan, 2002). If there was no difference in the scaling of H1 between a statement and question within an individual variety, there would be grounds for considering that the statement and the question had a unified source. This places us on slightly surer footing in relation to claims about similar scaling of H1 between Belfast statements and Belfast lists than between Glasgow statements and Belfast statements. However, in within-variety comparisons of scaling, scaling differences are not necessarily indicative of different historical origins either (see section 5.2.10 below).

The historical segmental phonetic distance studies assume the Neogrammarian model of language change. I raised the possibility that intonational change may be fundamentally different, involving aspects of pragmatic/discourse transfer. This further obscures the distinction between ancestry and contact in intonation as it introduces something like within-variety contact and perhaps analogical change e.g. the transfer of continuation intonation onto a statement context within a single variety. There may be some parallels here again with syntactic change (see Kroch, 2001). Pragmatic/discourse transfer in intonation may be somewhat like semantic change in the segmental lexicon but semantic shifts are problematic for existing phonetic distance measures because they create difficulties for the comparison of like with like. For example, German 'Hund' ('dog') needs to be compared against English 'hound', which is cognate with it despite the semantic shift, rather than with English 'dog' (Heggarty, 2000a, p.47). In intonation, I argue that Belfast nuclear statement intonation has arisen through transfer from another domain, probably list/continuation intonation. In the segmental lexicon, 'dog' has taken over the meaning originally encoded in 'Hund'/'hound'. 'Hound' now has a more specialised meaning. It is possible that the nuclear intonation contour for Belfast statements that was used before the transfer took place may have a more specialised function now.

5.2.3 Compatibility problem

In chapter 2, I raised the two problems identified by Heggarty (2000a) in attempting to measure phonetic distance: the Compatibility and Quantification problems. I regarded these as of central importance. My subsequent intonational work has convinced me that both are inescapable in intonation, despite the nature of the data and the measuring techniques being very different to those of segments. The Compatibility problem involves deciding which elements should be compared with which when making measurements. In segmental data, the

representation of the rhoticity contrast among English varieties is particularly tricky (Maguire, 2008, p.270). In the small-scale experimental work on segments that I presented in chapter 2, I showed that the **sdists** framework led me to compare post-vocalic /r/ in a rhotic variety against 'silence' in a non-rhotic variety and that this was a dubious procedure. I believe this shows the need for sound linguistic motivations in matching up elements for comparison.

Heggarty (2000a) determines like with like by creating a proto-form for each cognate and breaking this down into slots. The IPA transcriptions of cognate realisations in each of the varieties under study are also broken down into slots and matched up with each other with reference to the proto-form. This is feasible and robust when dealing with well-documented sub-families within Indo-European. In intonation, there is no such possibility of a proto-form. Like with like here requires the existence of a well-defined unit for intonational contrasts, the nuclear tone (see chapter 3). It is possible to match up $f0$ minima and maxima between utterances and between varieties using the nuclear tone as a template. For Heggarty, the concept of like with like means that the elements being compared were once the same at an earlier historical stage. In intonation, the Alignment hypothesis also posited that there was just one original alignment point for L1 and another for H1 historically. Note though the important distinction between segments and intonation here. Researchers match up segments between two realisations of a given cognate because they assume that in that position in the word, these segments were originally the 'same' source segment. I matched $f0$ minima and maxima in intonation because I wished to examine whether the corresponding minima and maxima could have had the 'same' origin respectively. Therefore, the cognate and the nuclear tone as frameworks have some comparability but there are striking differences too.

In segmental data, within an individual word, the feature values of a segment may be affected by surrounding segments and by other phenomena (e.g. low vowels typically longer than high vowels). It is important to be aware of this because it impacts on matching up like with like. I discuss this issue further in relation to segments and intonation in section 5.4.3 below. It is also relevant to the question of the independence of feature and parameter values (section 5.2.5 below).

5.2.4 Quantification problem

The Quantification problem deals with the explicit procedures involved in measuring and with exactly which features/parameters are measured. In Chapter 2, I argued that much phonetic detail incorporated into segmental phonetic distance measures fails to make an impact on overall scores. Phonetic distances could not avoid being affected by non-gradient phenomena like insertions of segments in IPA-based frameworks. I favoured a linguistically motivated but more sparse approach (Kessler, 2001, 2005; Kessler and Lehtonen, 2006; Kessler, 2007), choosing a few features that would be directly relevant to the researcher's hypothesis.

I have a similar conclusion about the approach to intonational measurements. For example, if my argument for the strong connection between Belfast statements and lists is a valid one, then it is really the H1 scaling measurements that are needed, not so much the other ones. Kessler (2001) demonstrated how irrelevant phonetic detail could obscure well-established connections actually being found by the measure.

Another question to ask is whether the measurements are actually accurate (Maguire, 2008, p.262). In relation to segmental distances, Heggarty (2006) severely criticises Edit Distances. Yet Maguire also raises the question of accuracy with Heggarty's own method, with the rejoinder that "phonetic symbols do not come with numbers attached". I would recast the problem somewhat differently to Maguire. Since phonetic symbols are not numeric, there are no fixed accurate numeric correspondences to them. It may not even make sense to think about accuracy of phonetic distances apart from the researcher's hypotheses.[2] In intonation, the continuous measurements of alignment and scaling that I made have not been numeric conversions of symbolic units. Yet without appropriate linguistic understanding, they have the potential to give misleading information about the particular hypotheses that interested me. So the issue of accuracy of measurements gives way to the question of how useful phonetic distance measurements actually are in evaluating hypotheses. A phonetic distance measure is useful if it successfully identifies cognates, if that was the researcher's goal. Likewise, measurements are useful if they show clear differences in intonation between sentence types, which help us to argue which two sentence types are more closely connected e.g. statements and lists.

---

[2]except in the trivial issue of whether the researcher has derived the correct numerical values from the procedure they claim to use.

Quantifying segmental phonetic distance often follows two prominent principles (Maguire, 2008, p.263):

1. avoiding selecting out a few special features in advance;
2. producing overall % distance scores between pairs of varieties (see also Nerbonne and Heeringa, 2010).

I found both of these expectations to be difficult in intonation. To begin with, the number of parameters available to me while assuming an AM framework was much smaller than the number of possible segmental features. Further, in order to evaluate my hypotheses I found it necessary to prioritise one parameter over another (e.g. prioritising the H1 alignment difference between Belfast and Glasgow English). If overall scores may be problematic, then pairwise comparisons will not follow easily. Pairwise comparisons are not as relevant in the context of the extremely small number of varieties I used for intonation as they are to the majority of segmental phonetic distance studies. I also considered just one phenomenon: statement nuclear intonation in Belfast and Glasgow English. By contrast, in segmental data, several phenomena will be embedded within phonetic transcriptions of a set of cognates. Further, my interest in distances between varieties was different to the prevailing interest in segmental studies. Segmental studies are interested in Language A being closer to Language B than to Language C as confirmation that Language A and B are more closely related than Language A is to Language C (I could easily recast this for the goal of cognate identification) (see chapter 2 for references). In intonation, Variety/Sentence type A being closer to Variety/Sentence type B than to C was relevant to whether both A and B had undergone the same specific intonational change and/or whether B might be derived from A (or vice versa).

Dialectometrical studies have the goal of deriving distances between varieties across several features. This is clearly different to my intonational approach of just using a few parameters. Although influenced by Dialectometry's prominence in the development of phonetic distances, my goal in intonation was more about exploring intonational change than simply deriving distances between varieties. Alternative approaches to the comparison of intonation would be possible, though. I could compare normalised $f0$ contours of the entire nuclear tone from Belfast and Glasgow English. This could use polynomials to capture the shape of the contours and express them using a set of coefficients (Grabe et al., 2007) or the Tilt model of Taylor (2000) (see section 5.3.7 below). This would

quantify differences between Belfast and Glasgow English. However, the co-efficients would not provide a transparent way of establishing whether differences between the 'Short' and 'Long' tails fit the predicted pattern of intonational change outlined in chapters 3 and 4. The polynomial approach could also be used in conjunction with normalised *f*0 alignment values (see Grabe et al. (2007, Appendix C)). This could only be done if the values were taken from utterances with identical text. Otherwise, the researcher may not be comparing like with like (see section 5.4.3 below). This approach would be a valuable addition to the individual alignment and scaling measurements I made in chapter 4. Yet they would best serve the purposes of modelling Belfast and Glasgow intonation for speech synthesis and quantifying their differences without clearly addressing the trajectory of intonational change.

The rates of language change and time depth also affect how appropriately distance measures may be used to study historical connections between varieties. In segmental data, when studying contemporary varieties of English, vowels are much more likely to display variability. Not enough time may have passed for any degree of consonant variability. Maguire (2008, p.262, 272) raises the point that rapidly changing phenomena e.g. loss of rhoticity, may obscure historical connections between varieties when using phonetic measurements. For this reason, Maguire casts some doubt on how reliable phonetic distance measures actually are as historical indicators. He suggests that his own method of distance based on lexical distributions of stressed vowel phonemes may be more suitable for studying historical links between varieties, precisely because it does not include a phonetic representation of rhoticity. At heart here is the avoidance of a feature known to be variable (see section 5.2.9).

In intonation, we have very little evidence of rates of change. HRTs were first noted separately in Australia and the U.S. around fifty years ago and today are reported in other parts of the English speaking world (McLemore, 1991; Sudbury, 2001; Shobbrook and House, 2003; McGregor, 2005). This would suggest that intonational change may be much faster than segmental change. Further, this particular change does not appear to be gradual phonetic change but rather the adoption of a particular contour in given discourse contexts. For phonetically based change in alignment, rates of change are unclear. Engstrand and Nyström (2002) argue for leftward alignment shift in the Swedish lexical accent over the course of c.50 years, drawing on data from the 1930s and 1970s. Hognestad (2006) proposes that the alignment of the peak in Norwegian accent I could drift from

early in the stressed syllable to well beyond it in less than 100 years. It is possible that the Belfast nuclear statement contour did emerge following the Alignment hypothesis. If so, the alignment of H1 has drifted so far that its path may have been obscured by the sheer extent of the change. Note though that it was not just the alignment difference between Belfast and Glasgow that led me to argue that they had separate origins but their within-variety behaviour. I agree with Maguire (2008, p.262) that phonetic distances are only one part of the picture in trying to address historical questions. They have an important place though, particularly in intonation, where at present there are so few alternatives.

Despite arguing against Heggarty on this point in chapter 2, I now believe that it is impossible to avoid phonology in relation to phonetic distances. The H1 alignment difference between the Glasgow and Belfast English data really reflects an underlying phonological distinction. The alignment scale may be broken down into intonational 'events' associated with the nuclear syllable (Glasgow) and 'events' associated with the IP-boundary (Belfast). Within the nuclear syllable 'event', there is important but smaller allophonic H1 alignment variability (Glasgow 'Short' and 'Long tails') and likewise in the IP-boundary 'event' (Belfast statements and questions). There is some kind of parallel here with Heggarty's approach to breaking down articulatory scales. Heggarty uses the number of phonemic distinctions typically made in place, manner and voicing respectively (Heggarty, 2000a,b; Heggarty et al., 2005). Therefore, Heggarty considers that major phonetic differences correspond to these broad phonemic distinctions and that there is lesser variability within these broad categories. This issue arises again in section 5.4.3 below.

I chose to measure the continuous acoustic parameters of alignment and scaling. The alignment parameter also required me to choose segmental landmarks from which to beginning measuring, e.g. nuclear vowel onset. I used these measurements on rises in the nuclear region only. This represents a small set of parameters, a small set of landmarks, and just one portion of the intonational contour of a complete utterance. I could have chosen to make other acoustic measurements, e.g. velocity and slope of the rise (see also Xu, 2005; Braun, 2006; Makarova, 2007).

My choice of parameters was motivated by their relevance within AM theory. Alignment is a fundamental parameter in making phonological distinctions in the AM framework (e.g. Prieto et al., 2005). There is also increasing evidence

of small phonetic differences between varieties (Ladd et al., 2009), and alignment has also been invoked in intonational/lexical accent/lexical tone change (see chapter 3). (Relative) Scaling is again fundamental in deciding which AM phonological tone to assign, e.g. H vs L vs !H (downstepped high) (see Ladd and Schepman (2003)). Along with *f*0 excursion, scaling is linked with the question/statement distinction. Although very little is known about historical change in scaling, knowledge of scaling changes in different prosodic contexts is increasing e.g. scaling effects of tonal crowding (Arvaniti et al., 1998, 2006; Arvaniti and Ladd, 2009).

Just as my intonational measurements represented a small subset of possible phonetic measurements, so too could segmental measures incorporate many more features/parameters than they actually do. There seem to be a few prominent motivations for the particular choice of elements: the dimensions and number of contrasts made by the IPA (Heggarty, 2000a; Heeringa, 2004) or other established feature system, the likelihood of change between segments (Baxter and Manaster-Ramer, 2000; Kessler, 2007) and computational ease (Heeringa, 2004; Beijering et al., 2008). Phonology is inescapable here too, as the IPA categories have been drawn up with respect to phonological/phonemic distinctions made in at least some languages (Ladefoged and Maddieson, 1996; IPA, 1999). In addition, it is impossible to think about diachronic segmental change without considering phonemic changes like mergers and splits. Again, this situation leads me to re-appraise my criticisms of Heggarty's reference to phonological distinctions in his own phonetic distance measures in chapter 2.

Straddling the segmental/prosodic divide, Pellowe et al. (1972) and Pellowe and Jones (1978) argued for the inclusion of as many features as possible. The reason for this is the possibility that by selecting just a few features, the researcher might end up discarding other features containing crucial information (Maguire, p.c.). However, if the researcher has a very specific hypothesis, then they may be in a position to choose the features/parameters most relevant to that hypothesis.

Of course there is a large difference between the continuous intonational measurements and discrete IPA categories. By using a feature such as [+/- voice], it has already been decided which elements should be [+voice] and which [-voice]. Even acoustic measures of segmental distance tend to rely on discrete bundles of spectral features per segment (Heeringa, 2004; Huckvale, 2004). With continuous measurements, these categories have not been set up in advance. Further, alignment measurements in intonation rely on the choice of segmental landmark

from which to begin measuring. There are not clear-cut guidelines on where alignment should be measured from, so differences between utterances, sentence types or varieties may emerge and disappear depending on the choice (Atterer and Ladd, 2004). In segmental data too, the fineness of transcribed detail may also have an impact e.g. the inclusion of diacritics will impact on whether two segments are transcribed the 'same' or not.

In chapter 2, one of my key questions was how much phonetic detail should be in a segmental measure and which features actually make a difference on overall scores and distinctions between varieties. I argued that it was really only large differences (that involved presence vs. absence of discrete segments) that made the impact. Gradations of phonetic distances did not make a difference to the groupings of varieties found. Including many features/parameters in a measure would be appropriate if these features actually contribute to the understanding of the patterns in the data. In intonation too, it is not really the fine gradations in the continuous measurements that are important in the underlying historical quest. It is bigger more discrete distinctions (e.g. between statistically significant differences and those that are not significant) (see chapter 4). There are the cases like the two Connaught Irish micro-dialects in which the small alignment differences may be perceptually salient (Dalton and Ní Chasaide, 2005). A measure with the goal of capturing this potential perceptual difference may need to focus on finer gradations.

Maguire et al. (2010), arguing against selecting out features in advance, have a valid point for their purpose of looking at overall patterns of convergence and divergence among English varieties. The continued emergence of rhoticity being the key issue dividing English varieties phonetically in these researchers' work suggests though that selectivity may be appropriate for some of their questions. The point is that the degree of fine-tuning of the measure and which parameters to include come down to the underlying questions of the study. A one-size-fits-all measure is inappropriate.

5.2.5 *Lack of independence*

In chapter 2, I briefly raised the little-discussed issue of phonetic features not being independent of each other. I also showed in chapter 3 how the intonational parameters of alignment and scaling, though theoretically independent of each other, in practice, often are not. In the Belfast and Glasgow English data, there

certainly were some dependencies between the parameters. For example, in the Belfast data, there was a trend for L1 and H1 to be earlier aligned in questions than in statements (see also Atterer and Ladd (2004, p.193) on Northern and Southern German). This indicates that the whole nuclear rise is aligned earlier in questions than statements.[3] The alignment and scaling of H1 did not appear to be independent, though actually in a different way to the findings of previous studies. Some studies report a correlation between a higher scaling of H and a later alignment of H in questions as opposed to statements (Gussenhoven, 2004; House, 2003; Makarova, 2007). However, in the Belfast data, questions indeed had a higher scaling of H1 than statements but an *earlier* alignment of H1. These findings give further support to my argument that a few well-motivated features in a measure may be most appropriate (see chapter 2, section 5.2.4 above).

5.2.6   Summary I

This section has illustrated how different quantification of phonetic distance among segments is to quantification in intonation in many respects (nature of measurements as continuous/discrete, nature of underlying questions, feasibility of an overall score). There are also some common concerns between them. One of these is the need for a few well-motivated features. A linguistically grounded approach does not have to mean extremely detailed in the manner of Heggarty (2000a). Another common factor is that a concept of underlying phonological categories may in fact be necessary in devising phonetic distance measures in segments and intonation and interpreting the results. Overall, I stress that the researcher's hypotheses should decide the approach to the quantification, the choice of features/parameters and how to measure them.

5.2.7   Presence or absence of historical data

Researchers of segmental phonetic data have the advantage of a large body of historical data, at least for well-documented language families like Indo-European. Knowledge of ancestral forms (though the Comparative method) is essential in the matching process of Heggarty (2000a). McMahon et al have also been able to produce phonetic transcriptions of historical dialects (McMahon,

---

[3]See also Asu (2006) who argues for examining the entire accent gesture, not just the peak.

2008).[4] Thus studies of segmental distance are now in a position to study historical and contemporary data side by side (see also Heeringa and Joseph, 2007). Intonation suffers from a lack of historical data even at very shallow time depths. Therefore, historical inference from contemporary data is the most feasible way of approaching historically-oriented questions. There is also the opportunity to consult archive recordings, even though the data would not be in a format particularly appropriate to intonational analysis, especially the phonetic measurements of alignment and scaling. I consulted the National Folklore Collection at the University College Dublin School of Irish, Celtic Studies, Irish Folklore and Linguistics. There I gained access to some recordings from speakers of Northern Irish English and Ulster Irish made in the 1940s, 1950s and 1960s. The typical Belfast nuclear rise of the contemporary data was not in overwhelming evidence in the archive data. There could be several reasons for this including:

- none of the speakers in the archive recordings were actually from Belfast;
- nuclear rises may not have been used as frequently in the region in the 1940s-1960s as they may be today;
- following from the above, the kind of transfer from continuation/list intonation to statement intonation that I argue for in relation to Belfast English may not have happened to a large extent at that time;
- style/register shift.

So I am unable to draw any firm conclusions from this exploration. Nevertheless, greater analysis of archive data would be an important direction for future work.[5]

5.2.8   *Apparent Time hypothesis*

Given the problems with historical data in intonation, another approach to studying intonational change would be based on the Apparent Time hypothesis. This is central to studies of Quantitative Sociolinguistics. This hypothesis states that synchronic differences between speakers of different ages also reflect the diachronic process of language change, i.e. that older speakers reflect more archaic

---

[4]They accept that such transcriptions are necessarily broader phonetic transcriptions than their transcriptions of data from contemporary varieties.

[5]However, I did observe nuclear rises in non-topic-final position and nuclear falls elsewhere in the speech of a few speakers. This was one factor leading me to consider the Transfer hypothesis as a viable account of the Belfast nuclear statement rise. Wells and Peppé (1996), who did impressionistic analysis of Belfast English data from the 1970s, did find the typical nuclear rising pattern.

patterns. Apparent Time and Real Time data have been compared in a key study by Bailey et al. (1991) lending support to the Apparent Time hypothesis. However, there were only 15 years separating the two sets of Real Time data. Therefore, it is unclear whether Apparent and Real Time data match up for considerably deeper time depths than this. Two classic earlier studies using Apparent Time are Labov (1963) on Martha's Vineyard and Labov (1966) on New York City English. In his Martha's Vineyard study, Labov compared differences between speakers of different ages with Real Time older linguistic atlas work and found an increase in the use of vowel forms localised to Martha's Vineyard. More recent work by Labov, e.g. Labov (2007) also uses the Apparent Time hypothesis to demonstrate the spread of the New York City short-/a/ system and the Northern Cities vowel chain shift. Piercy (2010), another study discussed in chapter 2 in relation to acoustic segmental distances, also uses Apparent and Real Time approaches to explore the stages in a phonemic split.

The Apparent Time hypothesis has rarely if ever been used in studies of intonation. In general, intonational studies have had little contact with the major quantitative sociolinguistic work done on segments. Yet an Apparent Time approach would be a feasible avenue for further work on Belfast and Glasgow English. In addition to the young adult speakers who took part in Experiment II in chapter 4, I could record much older adults from both of these varieties. I would predict that the older adults from the Glasgow variety would show earlier alignment of H1 in both the 'Short' and 'Long tails' conditions, in line with my proposal that there has been a rightward H1 drift in the few young adult Glasgow speakers reported on in chapter 4. It is more difficult to make predictions for the older Belfast English speakers. If the process of transfer of Question/Continuation intonation onto a statement context had not yet happened for them, I would predict that they would use nuclear falls consistently on statements or at least a much lower percentage of nuclear statement rises than the young adult speakers. Establishing the validity of such Pragmatic/Discourse-based change would be better served with a close study of intonation and discourse function in Belfast English than with an Apparent Time study, however. There is also the important issue of trying to ensure that in an Apparent Time study, the older and younger speakers would be matched along the other sociolinguistic dimensions than age. Otherwise, differences that emerged between them might not be due to the age difference and the process of change that represents but rather due to social class or others differences that might span all age groups. In chapter 4, I have already

acknowledged that I would need to focus much more on the sociolinguistic background of Belfast and Glasgow speakers for future work. For an Apparent Time study, the best approach might be to record young adults, their parents and perhaps also their grandparents, provided that all lived for the majority of their lives in the same region.

5.2.9 Stability and variability

In segmental phonetics/phonology, there is a wide body of knowledge about stable and variable features diachronically. In intonation, very little is known about stability and variability. Even within the synchronic domain, there are conflicting reports over which parameters are more stable than others. Xu (2005) (working on Mandarin lexical tone) found that alignment was more likely than scaling to remain constant. Arvaniti and Ladd (2009) found the reverse, the scaling mainly remains constant but alignment changes with tonal crowding in Greek data. Diachronic alignment and height changes have been reported in lexical tone in Chinese languages (Chen, 2000).

In chapter 3, I briefly discussed the stability and variability of intonational parameters in relation to previous studies. I now present an overview of the intonational parameters and how stable or variable they appeared to be in my data. My data would suggest the alignment of L1 is more stable than the alignment of H1 because it was unaffected by 'Tail' and by Variety. A common thread in previous intonational studies is that L target points are not as variable as H points (see chapter 3). However, the comparison of L1 and H1 alignment here is not balanced as I did not vary the number of preceding syllables before the nuclear syllable in the way that I varied the number of following syllables. Nolan and Farrar (1999) found earlier alignment of peaks with an increase in the number of preceding unstressed syllables.

Under the conditions in Experiment II, L1 scaling also appeared to be somewhat more stable than H1 scaling, as it was affected only by 'Tail' but again I did not set out specifically to study the relative stability of L1 and H1 scaling. As for whether alignment or scaling overall was more stable in my data, this is very unclear. Perhaps I could argue that H1 alignment is a bit more variable than H1 scaling, because there were significant differences in H1 alignment between 'Tail' types, sentence types and between varieties. There were significant differences in H1 scaling between Sentence types only. Even if this is a reliable assessment,

there is a further dimension which I have not yet considered. It is difficult to compare the magnitude of alignment and scaling differences. For a given amount of alignment difference, what is an equivalent amount of scaling difference? The few previous authors who do make suggestions about the stability and variability of alignment and scaling (e.g. Arvaniti and Ladd, 2009) do not address this issue (see chapter 2 section 8 and section 5.3.1 below).

Several studies show a consistently higher scaling of peaks in question intonation and/or under emphasis (Liberman and Pierrehumbert, 1984; Haan, 2002; House, 2003; Gussenhoven, 2004; Chen and Gussenhoven, 2008). There is much less consistency in the behaviour of valleys in these conditions (Pierrehumbert, 1980; Liberman and Pierrehumbert, 1984; Gussenhoven and Rietveld, 2000; Arvaniti and Garding, 2007). The recurrent findings on peaks may show that raising peaks is more likely than lowering but this is not certain.

5.2.10 Language vs. paralanguage

Separately, it would be wrong to ignore the paralinguistic dimension to intonation and this may be reflected particularly in the parameter of scaling. Are the concepts of stability and variability appropriate in a paralinguistic context? The relationship between language and paralanguage in intonation is not as clear as in segments. Testament to this are Gussenhoven's three Biological codes (Gussenhoven, 2004, ch.5), which he argues represent underlying paralinguistic tendencies in the use of pitch which have become phonologised (see also McMahon, 2007*a*). Two of these codes are particularly relevant to my intonational hypotheses: the Frequency code, which accounts for higher pitch in questions and the Production code, which accounts for the presence of rises in continuation contexts. Rises have traditionally been more associated with questions and continuation contexts than with statements and these have been considered to represent intonational universals (Bolinger, 1978). If intonation has a tendency towards more quasi-universal phenomena than segments (McMahon, 2007*a*), it confirms that it is much more difficult then to extract historical connections between varieties from natural independent developments.

Paralinguistic phenomena may be inherently gradient and bear a non-arbitrary relationship between sound and meaning. The primary consequence of this is that paralinguistic phenomena may not undergo sound changes in the way that

segments change phonetically. Rather they might ironically be stable in their gradience e.g. higher pitch might signal greater submission (question intonation) or greater non-finality (continuation intonation) at any time depth. It is also possible that change would happen to substitute a different gradient scale e.g. substituting more rightward alignment of H for higher pitch to express the same gradient meanings. This kind of change is very different to sound change which (though it may be phonetically gradual) results in the formation or collapse of discrete phonological categories (e.g. splits and mergers).

In chapter 4, I pointed out that without perceptual studies, it is difficult to assess whether the scaling differences between statements and questions in the Belfast and Glasgow English data represent paralinguistic gradient variability or whether they reflect a categorical phonological difference between a H tone in statements and a raised H tone in questions. Even within existing perceptual studies, there are conflicting results about the behaviour of the scaling of intonational tonal targets. Though Remijsen and van Heuven (1999) found categorical perception between H and L targets in a specific context, other studies indicate that scaling is gradiently variable in perception (Ladd and Morton, 1997).

5.2.11 Intonation and articulation

The study of the coordination of intonation and articulatory gestures has the potential to lead to a greater understanding of stability. From personal communication with Doris Mücke on this topic, the in-phase mode of coordinating tonal and segmental articulation is the "most stable (intrinsic) mode in gestural timing" (see also Mücke et al., to appear). This would lead to tonal targets being aligned acoustically around the beginning of the syllable or slightly before it. Indeed a recurring finding cross-linguistically is for the L of rises to occur around the beginning of the nuclear syllable or just before it (Prieto et al., 1995; Arvaniti et al., 1998; Ladd et al., 1999, 2000; Ladd and Schepman, 2003).

In Experiment I, the low turning point of the Belfast rise was consistently aligned around the end of the nuclear vowel/syllable. In Experiment II, L1 in the Belfast data also showed a strong trend of being aligned towards the end of the nuclear syllable. The L in Belfast rises could be in-phase with the consonant and vowel articulation at the beginning of the postnuclear unstressed syllable (i.e. aligned simultaneously with the beginning of the consonant and/or vowel gestures). Another possibility is that the Belfast L may be aligned with the peak

velocity of the following consonant closure, as for the L in the French early rise, the H in Italian statements or Catalan nuclear rises (D'Imperio et al., 2007; Prieto et al., 2007).

As research on alignment and articulation continues, a small number of prominent articulatory gestures may emerge that constantly align with tonal targets. This would provide a further way of quantifying alignment change. In a synchronic context, Mücke et al. (2009) argue for an "anchor shift" between prenuclear and nuclear accents to account for the alignment difference between them. If a change in articulatory anchoring point is possible in a synchronic point, perhaps this might also be possible in a diachronic context. Mücke (p.c.) believes this is possible. The question of the relationship between the synchronic and diachronic in articulatory gestures in segments emerges in Browman and Goldstein (1989); McMahon et al. (1994); Blevins and Garrett (1998) and Blevins (2004).

5.2.12 Summary II

In this section, I have demonstrated the following important points in my comparison of segmental and intonational phonetic distance measurements:

- The nature of the historical questions I asked in relation to intonation was quite different to the typical historical questions underlying segmental measurements.
- The distinction between ancestry, contact and chance resemblance is much less clear cut in intonation.
- Having a concept of comparing like with like is as essential for intonation as it is for segments.
- I argue that for segments and for intonation it is crucial to choose a few features/parameters that are directly relevant to one's hypotheses.
- Absence of historical data is a problem for intonation but archive recordings offer possibilities for the future.
- Stability and variability are much harder to assess in intonation than in segments.

5.3 Overall measure of intonational distance and Phylogenetics

I used the parameters of alignment and scaling to compute intonational measurements but thus far have not described an overall measure of intonational

distance. In segments, phonetic distance measures are usually combinations of several feature values. Here I outline how the intonational parameters could be combined into an overall score. However, I believe that a combined measure of intonational distance is not really appropriate for studying intonational history. This is because it is much harder to rule out independent similarities (homoplasy) in intonation than in segments. It would be inappropriate to use the distance of two contours on a particular parameter as support for a hypothesis in which that parameter was irrelevant e.g. the similar scaling of L1 and H1 between the Glasgow and Belfast data is largely irrelevant to the Alignment hypothesis (see also section 5.2.2 above). I believe those devising segmental measures should consider this possibility in relation to segmental hypotheses too.

That said, an overall measure of intonational distance would have the boon of enabling researchers to compare the same set of varieties for their segmental distances and intonational distances respectively. Lohr (2000) compared lexical and phonetic distances on the same data. Nakhleh, Warnow, Ringe and Evans (2005) compared language trees computed using different combinations of their lexical, phonological and morphological Characters (see chapter 2). Thus there is interest in measuring the same data in different ways and examining to what extent the results correlate. Comparing segmental and intonational distances is therefore an avenue for future work.

One of the biggest problems with an overall measure of intonational distance would be how to weight the different parameters relative to each other. I discuss this in section 5.3.5 below. There are also difficulties for combining features/parameters into an overall score if these elements are not independent (Embleton, 2000) (see section 5.2.5 above). First, I explore how to combine the parameters in the absence of any weighting system.

5.3.1   Basic measure outlined

For each nuclear contour in my data, I have four different measurements:

1. L1 alignment
2. H1 alignment
3. L1 scaling
4. H1 scaling.

The first step is to recognise that alignment and scaling are measured on two different scales. Alignment is measured in milliseconds (ms). I computed a normalised scaling value for L1 and H1 in semitones. To combine alignment and scaling measurements into a single score, it would be necessary to express them both on the same scale. This requires that I discretise both continuous scales to create some kind of equivalence between 'units' of alignment difference and 'units' of scaling difference. The units on this scale could correspond to the Just Noticeable Difference (JND) for alignment and pitch height. Braun (2006, p.473) reports a JND for alignment of 50 ms and a JND for pitch height of 1.1 semitones (sts).[6] I use these in this first attempt.

5.3.2 *Poles of the distance scale*

The next step is to put ends on these scales i.e. to express the values between 0 and 1. Choosing these is a somewhat arbitrary matter. For example, I measured the alignment of L1 from the beginning of the nuclear vowel (Begv1). Which point should correspond to 0 on the new alignment scale? If I consider 0 to be Begv1 itself, then there is the problem of the occasions in which the alignment of L1 occurred before v1. I see no easy solution except to use the Belfast and Glasgow data themselves. I adopted a similar procedure in relation to the similar problem of trying to put ends on distance scales in segmental measures (see appendix B and chapter 2 section 7). Future work on intonation and articulation may identify the earliest possible gestural coordination point(s) for L1 alignment. This could be used as 0 on the alignment scale. Returning to my data, the earliest L1 occurs before v1 in an individual utterance is 243 ms before Begv1. So we may begin the L1 alignment scale at 250 ms before Begv1 and move incrementally in steps of 50 ms. The latest L1 occurs after v1 is 475 ms so the L1 alignment scale may end at 500 ms. So the L1 alignment scale will consist of 15 steps. In relation to H1 alignment, I measured it from the end of the nuclear vowel (Endv1). The earliest H1 was aligned was 157 ms before Endv1, so I may start the H1 alignment scale at 200 ms before Endv1, again moving in steps of 50 ms. The latest H1 occurred after Endv1 was 735 ms, so 750 ms is a reasonable place to end the H1 alignment scale. So the H1 alignment scale has 19 steps. It is not ideal that the two alignment scales do not contain the same number of steps each, as

---

[6]Even though I reported proportional alignment results in chapter 4, I prefer here to focus on absolute alignment measurements for the sake of simplicity. 't Hart et al. (1990, p.28-9) put the JND at 1.5-2 semitones.

in the present format H1 alignment is automatically weighted heavier than L1 alignment (see Kondrak (2000) in relation to segments in this regard).

The highest value for H scaling was 7.772 semitones above the speaker's mean. The lowest value for H scaling was 3.443 semitones below the speaker's mean. So I could set the scale for H1 scaling at 4 semitones below the speaker's mean, moving in steps of 1.1 sts until 8 semitones. This would have 10.909 steps. For simplicity, I could modify the 1.1 st JND threshold and assume that it is just 1 st. This would give 12 steps in this scale.

The highest value for L1 scaling was 2.035 semitones above the speaker's mean. The lowest value was 5.533 semitones below the speaker's mean. So I could set the L1 scaling scale at 6 semitones below the speaker's mean and move in steps of 1.1 semitones to 2.1 semitones above the speaker's mean. This would have 7.364 steps or 8 steps if I take the more straightforward value of 1 st to be the step size. Liberman and Pierrehumbert (1984, p.218) report from their experiments on prominence that Ls may be closer to the speaker's lower baseline than Hs are to the higher baseline of the speaker's range. This means it is not surprising that there are fewer steps in the L1 scaling scale than for the H1 scaling scale. It is not ideal that the scales for L1 and H1 scaling have fewer steps than the alignment scale as this again puts an implicitly heavier weighting on alignment. I emphasise the exploratory nature of combining these intonational measurements and I do not intend to draw strong conclusions from them.

5.3.3 Combining the scales

I now provide a few basic examples of how to combine the scales for alignment and scaling and what kind of results this might give for my data. The examples use the mean values for alignment and scaling from each of sixteen categories. All of these mean values may be found in appendix D. These categories are based on variety (Belfast, Glasgow), sentence type (4 types) and 'tail' status (2 'tail' types) (2 x 4 x 2).

For alignment, I work in steps of 50 ms. Therefore, if the mean alignment of L1 or H1 between two categories was less than 50 ms, there was a score of 0; if it was above 50 ms but within 100 ms, there was a score of 1, and so on. For scaling, I work in steps of 1 semitone. If the mean scaling of L1 or H1 between two categories was less than 1 semitone, there was a score of 0; if it was above 1 semitone but less than 2, there was a score of 1, and so on. Once I did pairwise

comparisons for all categories for each of the 4 parameters separately, I simply added the 4 scores per pairwise comparison together to get an overall score of intonational distance between each pair of categories. This is essentially applying the Manhattan distance (see chapter 2 and Gussenhoven and Rietveld (1991, p.442)). To get a score that falls within the range of 0 to 1, I could divide these scores by 54. This is the total number of steps if I combine the number of steps on the alignment and scaling scales (alignment L1 15 steps, alignment H1 19 steps, L1 scaling 8 steps, H1 scaling 12 steps).

The results revealed that the closest pairs of categories (all receiving a score of 0) were the following:

1. Belfast 'Short tails' statements and lists (along with the Glasgow short tails lists);
2. Belfast 'Long tails' wh and y/n questions, and
3. Glasgow 'Long tails' statements and wh questions.

The highest score was a score of 16 (16/54 = 0.3), scored between the following three pairs:

1. Glasgow 'Short tails' statements and Belfast 'Long tails' statements,
2. Glasgow 'Short tails' statements and Glasgow 'Long tails' lists,
3. Glasgow 'Short tails' wh questions and Glasgow 'Long tails' lists.

In relation to the Alignment hypothesis, the pair of Glasgow 'Short' and 'Long' statements scored 5. The Glasgow and Belfast 'Short' statements scored 8. The Glasgow 'Long' statements and Belfast 'Short' statements scored just 3. The Belfast 'Short' and 'Long' statements scored 4.[7] It is difficult to make assessments about "Re-alignment" on the basis of a combined score and these results do not clearly show the important distinction between Belfast and Glasgow.

In relation to the Transfer hypothesis, the Belfast 'Short' statements and lists are closer to each other than the statements are to either of the question types. The same pattern is true of the 'Long tails' data. The Glasgow 'Short' statements are closer to the Glasgow 'Short' wh questions, next to the 'Short' y/n questions and

---

[7]It is probably inappropriate to compare the Belfast 'Short' and 'Long tails' statements because there is a problem with absolute alignment from Endv1 in this variety. Absolute measurements show a large difference in alignment between the 'Short' and 'Long tails' sets but this is a consequence of there simply being more unstressed syllables in the 'Long tails'. H1 is usually aligned in the IP-final syllable in both the 'Short' and 'Long tails' (see chapter 4).

then to the 'Short' lists. As with the Belfast data, this pattern is repeated in the 'Long tails' data. This supports my argument that the Belfast statements may have derived from Belfast list intonation rather than question intonation. It also supports my argument that Glasgow statements and wh questions may have a special connection.

I consider that these combined distance scores at least partially reflect the conclusions I have drawn separately from the examination of statistically significant differences in measurements for the parameters individually. This is somewhat surprising given that the parameters had different functions in evaluating the different hypotheses (e.g. scaling is most relevant to the Transfer hypothesis and H1 alignment is most relevant to the Alignment hypothesis).

However, the patterns found by the combined score might also be found by simply using distances from an individual parameter. Further, this combined measure is extremely simplistic at present and there are several issues I have failed to address (see section 5.3.5 below). What is heartening is that it is possible to develop some kind of measure in a similar way to the development of measures of phonetic distance in segments (e.g. using Manhattan distances). This shows that all the work that has gone into developing these measures has not been completely in vain for applicability to new domains like intonation. The distance scores for each of the parameters as well as the complete distance matrix may be found in appendix E.

5.3.4 *NeighborNet*

I used the overall distance matrix as the input to a NeighborNet network (Huson and Bryant, 2006) in figure 5.1.

This gives a visual representation of the key points from section 5.3.3 above. The network also demonstrates that the within-variety variability is generally less than the between-variety variability. This would be of interest to those who primarily work on segmental distances between varieties and languages.

5.3.5 *Weighting*

The previous section has outlined one way of combining the parameters of alignment and scaling into an overall measure of intonational distance. It was a very simple outline and deliberately did not apply special weightings to the different
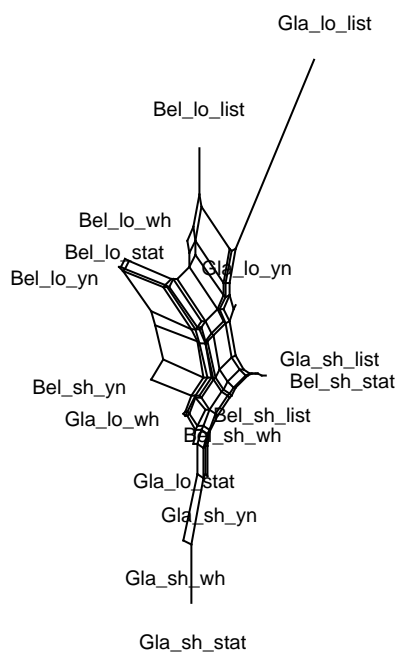
Figure 5.1: NeighborNet network for the combined intonational distance scores. Notice that most of the Belfast categories are located at the top left of the network diagram and that most the Glasgow categories are located at the bottom right. Abbreviations: Bel (Belfast English), Gla (Glasgow English), lo ('Long tails'), sh ('Short tails'), stat (statement), wh (wh question), y/n (y/n question)

parameters. In chapter 2, I suggested that distance methods might offer intonation the possibility of hypothesis testing and exploration. Trying to find a consistent weighting system for these intonational parameters would appear to be extremely difficult. What constitutes appropriate weighting for segmental features in segmental phonetic distances may also actually be a less clear-cut issue than researchers acknowledge. For example, some early work of Nerbonne and colleagues (Nerbonne and Heeringa, 1997) experimented with different weightings but these weightings did not give better results. So for Nerbonne & Heeringa, successful weighting means weighting in such a way as to produce results which correlate well with an external assessment of the 'right' outcome (in their case, dialectologists' judgements). For Heggarty (2000a) by contrast, weightings are linguistically motivated using criteria such as phonological behaviour in typological patterns and phonemic distinctions. In chapter 2 I was critical of Heggarty's recourse to phonemic distinctions for setting weightings in what he explicitly claimed was a measure of phonetic distance. My intonational work has given me a modified attitude though about the feasibility of divorcing phonological behaviour from phonetic measurements (section 5.2.4 above). Huckvale

(2004) does not discuss the issue of weighting parameters in his acoustic measure of segmental phonetic distance. Heeringa (2004, p.109-10) only appears to apply weightings in his acoustic segmental measures in relation to duration of segments.

Nerbonne & Heeringa's approach to weighting would be unsuitable for my intonational questions. I queried specifically whether the nuclear statement contour in Belfast and Glasgow English did indeed represent the 'same' phenomenon. Therefore, it would not make sense to weight the parameters in such a way as to ensure that there would only be a small distance between the Belfast nuclear statement contour and the Glasgow nuclear statement contour. Instead I try to examine whether there are sound linguistic motivations for weighting parameters more heavily than others. I also admit I explore weighting simply based on my own data set, how the patterns in the data seem to relate to the hypotheses and the difference between inter- and intra-variety comparison in terms of precedents about stability and variability (section 5.2.9 above).

In relation to the Transfer hypothesis, the most important parameter would appear to be H1 scaling. Changes in pitch height are perceptually salient (Makarova, 2007) (as are pitch range changes (Gussenhoven and Rietveld, 1991, p.445)) and could be one motivation for weighting H1 scaling more heavily than any other parameter in my small set (see also Grice, 1995; Arvaniti et al., 2000). This clashes though with the lack of knowledge about how scaling behaves at a cross-varietal level instead of between sentence types in a single variety. It is plausible to use the H1 scaling difference between Belfast statements and questions as a 'salient' difference. However, it is more tenuous to suggest that the absence of a difference in H1 scaling between the Belfast and Glasgow data entails that they must be perceived very similarly.

If I just compare the two scaling parameters, there would be arguments for weighting H1 scaling more heavily than L1 scaling, given the lack of consistency in the behaviour of valley scaling in previous studies (section 5.2.9).

Weighting H1 scaling heaviest would create problems for the overall point of my intonational analysis: that Belfast and Glasgow English nuclear intonation contours are fundamentally different to each other and have separate origins. The key difference between them was in the alignment of H1. If scaling differences represent a greater amount of distance than alignment differences, then there would be scope to claim that the Belfast and Glasgow nuclear statement contours

are actually very similar. There would be some grounds for trying to claim that a difference in the alignment of a peak (Belfast vs. Glasgow difference) is greater perceptually than a difference in the alignment of a valley (the larger alignment difference in the case of Belfast statements vs. lists) (Dilley, 2005) but this does not solve the problem here. On the other side, my data seemed to show L1 alignment as more stable overall than H1 alignment. This might suggest that an L1 alignment change should be weighted heavier than a H1 alignment change.

The conclusions I draw from this exploration of weighting intonational parameters against each other include that in the absence of more perceptual tests, it is difficult to come up with absolute assessments of whether alignment or scaling measurements should be weighted more heavily than the other. In addition, it may be impossible to have objective weightings external to the individual datasets and hypotheses in intonation. Given these problems, perhaps weighting in intonation should be avoided but this leaves me open to the charge of combining parameters too crudely.

5.3.6   *Which parameters to include/exclude*

Apart from the problem of weighting different parameters, there is also the question of whether I should exclude any of the parameters altogether. Certain researchers of segmental distance have eschewed features known to be variable (section 5.2.4 above). Of my set of intonational parameters, the one parameter with the most indication of being variable over time is the alignment of H1. Excluding it would not be satisfactory for two main reasons. First, the crucial link with the Alignment hypothesis would be immediately lost. Second, as I have emphasised elsewhere (e.g. section 5.2.9 above), even if the nuclear tone contours from two varieties were very similar along the other parameters, there are few grounds for arguing that these similarities could not have come about by chance or independent developments. Researchers on segmental distances may exclude variable parameters because their underlying questions are not about the actual paths of sound change. Their questions are about confirming and discerning ancestral connections between languages. By contrast, I was specifically interested in documenting potential paths of intonational change as displayed in patterns in synchronic data. In addition, once segmental researchers exclude variable features, they have a surer foundation for what they consider to be stable features than is possible at present for intonation.

There might separately be arguments for omitting the parameter of L1 scaling from a combined measure for my data, because there were no differences of Sentence type or Variety in the Sentence reading task along this parameter.

A combined intonational measure might be relevant in modelling Belfast and Glasgow contours for speech synthesis purposes and perhaps the combined scores would correlate with perceptual distance judgements. With an expanded study of more varieties, it might help to explain why these two varieties might be considered similar, regardless of whether those similarities reflect any historical connection (i.e. contact in the case of intonation, see section 5.2.2 above) or not. So there are potential values of some kind of combined distance measurements in intonation. What I emphasise at present though is they do not clearly address historical questions in intonation and without perceptual or other studies, it is unclear what exactly these combined distance scores represent linguistically.

If there are breakthroughs in the understanding of which elements of intonation change over time and which do not, it would be possible to revisit combined intonational distance scores and examine if they provide decent proxies for intonational change or not. After all, segmental distance measures that researchers have designed to be proxies for historical issues such as cognate identification and reconstruction have done so only on long-standing pillars of existing knowledge, particularly through the Comparative method.

5.3.7 *Taylor (2000)*

I have just proposed a way to combine the intonational parameters of alignment and scaling into an overall measure of intonational distance and have demonstrated some resulting problems. An existing measure of intonational distance from the field of speech synthesis is the Tilt measure within the Tilt model of Taylor (1992, 2000). Essentially, Taylor's measure is a measure of contour shape of intonational 'events'. Specifically, it involves comparing the amount of rise and fall in an event.

$$\text{tilt} = (\frac{|A_{\text{rise}}| - |A_{\text{fall}}|}{2(|A_{\text{rise}}| + |A_{\text{fall}}|)} + \frac{D_{\text{rise}} - D_{\text{fall}}}{2(D_{\text{rise}} + D_{\text{fall}})})$$

(Taylor, 2000, p.1704)

In that intonational event, the Tilt measure takes into account the Amplitude (A) of the rise and fall elements along with the duration (D) of the rise and fall elements. By Amplitude, Taylor here means the $f0$ excursion.

I provide a brief outline of how Taylor's measure could be applied to my Belfast and Glasgow English data. I begin with Belfast statements and treat the intonational 'event' as the nuclear tone. I assume that there is a rise and no fall within the nuclear tone because the Belfast utterances did not contain a final low. To work out the Amplitude of the rise, I prefer to use semitone values from the normalised scaling parameters rather than the speaker dependent linear Hz $f0$ excursion values. The mean normalised L1 scaling value for the Belfast statements 'Long tails' was 2.274 semitones below the speaker's weighted average. For H1 scaling, it was 1.808 semitones above the speaker's weighted average. Adding these together gives the $f0$ excursion of 4.082 semitones.

$$(4.082 - 0)/2(4.082 + 0) = 0.5$$

This is the Amplitude part of the measure.

To work out the duration of the rise, I use the alignment parameter, expressing L1 and H1 with respect to the same segmental landmark. In appendix D, there are mean values for the alignment of L1 with respect to the end of the nuclear vowel (Endv1). In the 'Long tails', the mean absolute alignment for the Belfast statements was 15 ms before Endv1. The mean absolute alignment of H1 with respect to Endv1 was 458 ms after Endv1. The duration between L1 and H1 is then 458 + 15 = 473 ms.

$$(473 - 0)/2(473 + 0) = 0.5$$

This is the duration part of the measure. Adding the Amplitude and Duration parts together gives a score of 1 for the Belfast statement rise.

A plain rise will always receive a score of 1 in Taylor's measure. Therefore, it would not expose the difference in H1 scaling between statements/lists and questions that I found in the Belfast English data in particular. Although Taylor's measure takes the excursion into account, it can only contrast the excursion

(Amplitude) of a rise with that of a fall, not of two different rises. Therefore, Taylor's measure would probably assign the same score or almost the same score to all Belfast sentence types. Of course, a number of the Belfast questions had small slumps after the main nuclear rise. I did not measure these specifically, but a slump with a small $f0$ excursion and a short duration would not impact much on the score of 1. It would reduce it only slightly (see Taylor, 2000, p. 1705-6).

For the Glasgow statements, I run into the problem that I did not make an exact measurement of the final low after the nuclear rise (see chapter 4). Nevertheless, I can make an approximation of the fall component for the purposes of this preliminary exploration.

The mean alignment of L1 in the 'Long tails' was 80 ms before Endv1. The equivalent value for H1 was 125 ms after Endv1. The duration of the rise component is therefore 205 ms (125 + 80). The mean scaling of L1 was 1.202 semitones below the speaker's weighted average. For H1, it was 1.242 semitones above the speaker's weighted average. The Amplitude of the rise component is thus 2.444 semitones (1.202 + 1.242).

I make the simplifying assumption that the scaling of the final low is the same as that for L1 (actually, it was usually lower than the scaling of L1, reflecting the reported pattern of a fall to the bottom of the speaker's range (Ladd, 2008, p.127)). This is because I did not have an exact measure of the final low in the Glasgow data (see chapter 4 section 4.11.2). Therefore, the Amplitude of the fall component would be the same as for the rise component.

I also make the simplifying assumption that the final low is aligned at the very end of the utterance. I have time measurements for Begv1 and for the utterance end so can work out a mean distance between the two for the Glasgow statement data. The mean distance between Endv1 and the utterance end was 575 ms. Therefore, the distance between H1 and the utterance end is 575-125 = 450 ms. This is the duration of the fall part of the Glasgow nuclear statement contour.

Amplitude portion:

$$2.444 - 2.444/2(2.444 + 2.444) = 0$$

Duration portion:

$$(205 - 575/2(205 + 575)) = -0.2371795$$

The overall Tilt value for the Glasgow 'Long tails' statements is thus -0.2371795. This captures a contour shape that has both a rise and a fall with more of a fall than a rise. A plain fall would always score -1 and a contour with equal amounts of rise and fall would score 0 (Taylor, 2000, p.1705). Though crudely computed, this gives support to my argument that the Glasgow statements represent "Re-alignments" of nuclear falls.

So Taylor's measure is useful when dealing with different contour shapes because this is precisely what Tilt measures. It cannot account for differences in magnitude/pitch range of contours which share the same overall shape and this is problematic for the Belfast data. The fact that all Belfast sentence types would receive the same or nearly the same Tilt score does reflect that some kind of transfer is plausible for the Belfast statements, however. This section has outlined how an existing combined measure of intonational distance might be used with my data and this is valuable in itself.

5.3.8 *Characters and continuous measurements*

In chapter 2, I made the broad distinction between Distance- and Character-based phylogenetic methods. I also pointed out that it would be extremely difficult to incorporate continuous intonational measurements into a set of Characters given the current set up of Character methods like 'Network' (Bandelt et al., 1999).[8]

Surprisingly though, I realised that my approach to evaluating the intonational hypotheses had connections with the computational cladistic[9] method of Nakhleh, Warnow, Ringe and Evans (2005), which uses discrete Characters. So although I do not believe that my data could be recast in a format enabling me to use their methods with intonation, it is worth exploring some of these connections in overall approach. Nakhleh et al.'s method is not a method of measuring

---

[8]Characters do not have to have discrete values in the biological uses of these methods but at present, it is extremely difficult to use these methods with data in any other format than biological data.

[9]This term, originally from Evolutionary biology, refers to an approach that aims to study only those connections between languages that have come about through ancestry and the principle of Descent with modification i.e. not borrowing or other connections between languages.

phonetic distance. It contains three elements: lexical, phonological and morphological. Lexical Characters involve cognate judgements. The phonological and morphological Characters represent complex series of sound changes. The Character states are essentially binary, i.e. whether two languages are cognate for that item or not, or whether they have undergone that sound change or not.

In evaluating the Alignment and Transfer hypotheses, I prioritised individual parameters over others, e.g. the alignment H1 in relation to the Alignment hypothesis. In the very different context of Indo-European reconstructions in Nakhleh et al.'s work (Nakhleh, Warnow, Ringe and Evans, 2005, p.188), they too feel the need to select out an extremely small number of Characters (just 4 from their set of 336) to act as major arbiters in the choice of one Reconstruction method over another.

Their justification for their selectiveness is that the Characters they have chosen represent the most robust Characters, most indicative of ancestral connections between Indo-European languages and thus least likely to be affected by borrowing or homoplasy. As I have made it clear in this chapter, it is much less clear in intonation which parameters are more robust than others. In addition, the concept of robustness may be different between segments and intonation. In segments, robustness may be defined empirically using existing findings about historical stability. In intonation, robustness may be more functional, e.g. which parameters are actually used to distinguish varieties and sentence types, regardless of their historical stability.

There are some parallels in how Nakhleh et al. use the few important Characters in their data analysis. Of the 4 crucial Characters, Nakhleh et al. use different groupings of them to argue against Indo-European trees found by various reconstruction methods. Some Characters support certain groupings within the trees; other Characters support other groupings. For example, P1 (phonological Character 1) requires the Italic and Celtic branches to be closest to each other in the form of being 'sisters'. However, M6 and M8 (morphological Characters 6 and 8) allow for Italic and Celtic to be nested one inside the other and not necessarily 'sisters' as long as there is no overlap between Italic and Celtic and the separate part of the tree with Germanic, Greek and Indo-Iranian. Nakhleh, Warnow, Ringe and Evans (2005, p.190) point out that a preference for a particular tree depends crucially on which Characters the researchers think are most important.

Nakhleh et al.'s Character set and Character states are not universal but rather very specific to their Indo-European data. There are parallels here too with my intonational work as my application of the alignment and scaling measurements was specific to the Belfast and Glasgow English data and could not be replicated straightforwardly elsewhere.

Nakhleh, Warnow, Ringe and Evans (2005) conclude that linguistic expertise is essential in the context of quantitative methods. I concur entirely in relation to intonation as well as in relation to their own work. I stress that linguistic expertise is not an excuse for a lack of objectivity. It makes sense to use the knowledge available of which linguistic elements are more likely to change than others and in which directions to attempt quantification in the context of historical hypotheses.

Of course, the kinds of categories used by Nakhleh et al. are very different to the intonational categories I used. Their categories are languages represented by a set of words and by phonological and morphological features. My categories are individual sentence types compared both within- and between- varieties represented by a small set of intonational parameters. Still, in some respects, my intonational work has more in common with Nakhleh et al.'s work on Characters than on existing work on segmental phonetic distances.

5.4 What can intonational distances offer segmental phonetics?

This chapter has demonstrated some stark differences in the nature of the intonational measurements I made and existing measures of segmental phonetic distance. In chapter 2, I addressed which elements of segmental phonetic distance measures could be relevant to measuring intonational distances. Now it is time to take an alternative perspective. Instead of using segmental measures as the baseline against which to compare the intonational approach, I ask what benefits the exploration of intonation could bring to approaches to phonetic distance in segments. Segmental phonetic distance measures have remained largely restricted to isolated word comparison. My attempt at intonational distances has been a first step away from this.

5.4.1 Multiple renditions vs. single transcribed words

A major difference between approaches to measuring phonetic distance between segments and intonation is in the nature of the data. Articulatory segmental distances have commonly been applied to isolated transcribed words. In intonation, I used acoustic data of complete utterances. Articulatory segmental distance approaches to date have hardly ever incorporated within-variety variability, though McMahon et al. (2007) believe it is the next step. Such variability underpinned my use of acoustic intonational data. Acoustic segmental distances have been used in conjunction with within-variety variability, however.

Transcribed segmental data of course has its problems, including intertransriber reliability (Maguire, 2008, p.261-2) and the more basic question of whether the use of phonetic symbols is phonetically faithful in the first place (Ladd, forthcoming). This immediately suggests problems for Edit Distances which rely precisely on such discrete symbolic data.

The study of intonational measurements in the context of potential historical change suggests that segmental approaches may be missing out on information about historical change by restricting themselves to a single token of each transcribed word. It was precisely the variability in the alignment of H1 between the Glasgow 'Short' and 'Long tails' conditions that allowed me to argue for the Alignment hypothesis as an account of the Glasgow nuclear statement contour. There would be an argument that this synchronic variability does not lead to a direct inference of historical change but merely displays the presence of stable synchronic variation between utterances with differing numbers of syllables in the 'tail'. Nevertheless, if researchers in the segmental field wish to continue to understand which varieties/languages are more retentive than others, then synchronic variability is one way in which to expand this understanding. This would involve calculating within-variety distance scores that could then be compared across varieties. Such an approach has been carried out for in a few acoustic segmental distances (Huckvale, 2004, 2007; Ferragne and Pellegrino, 2010).

This could examine the phenomena of vowel shifts. Although these have been examined in acoustic segmental distance studies, there are still several unexplored avenues. Using allophonic variation in individual varieties, acoustic segmental distances computed from $f1$ and $f2$ scores and vowel durations could examine a phenomenon like the Modern English Vowel Shift Rule (McMahon,

2007*b*). Here there are two types of alternations: CiV Tensing (e.g. various, variety) and Trisyllabic Laxing (e.g. divine, divinity). The alternations are morphophonologically conditioned and may be part of a chain shift. There is somewhat parallel to the prosodic conditioning (number of unstressed syllables following the nuclear syllable) which accounted for the H1 alignment variability in the Glasgow 'Short' and 'Long tails'. Different English varieties may have undergone these vowel shifts in subtly different degrees. It may be most appropriate to think of this as an examination that could be carried out, were the Modern English Vowel Shift still a change in progress. A comparison of the $f1$ and $f2$ values and vowel durations in the appropriate vowel in each alternation within each variety individually would be the first step. This would show the relative degree of Tensing or Laxing respectively. Then these relative distances could be compared across varieties showing which varieties had undergone the most Tensing or Laxing respectively. Using within-variety distances echoes the within-language distance scores of Ellison and Kirby (2006) computed before any cross-linguistic comparison takes place.

Examining within-variety variability could also be used to examine the directions of vowel changes in chain shifts (e.g. Northern Cities shift), posited in the set of principles by Labov (1994). Though studies such as Torgersen and Kerswill (2004) have used a number of varieties of English and other languages as case studies for testing Labov's principles, there are still new avenues for quantifying the degree to which a given vowel in a variety follows Labov's principles or not. For example, the directions of change are different for long and short vowels respectively.

A brief consonant example would be the following: if variety A had an alternation between [k] and [kʲ] but variety B only had [k], then there might be signs of a change, but studies of segmental distances have not focused on this, with the possible exception of Maguire et al. (2010).

The intonational measurements have shown the potential value in opening up segmental quantification to measuring degrees of sound change and not just overall distances between varieties/languages. Intonation admittedly is as yet very much tied to synchronic variation and to the attempts to extrapolate back from this. Yet many current segmental distance studies have not exploited the rich body of evidence on synchronic variation that is open to them.

5.4.2   *Different types of linguistic change*

Existing work on segmental phonetic distances has been far more concerned with the connections these distances find between varieties and languages than with actually quantifying the types of sound changes that may separate and connect these varieties and languages. In intonation, I demonstrated the possibility that there may be two main mechanisms of change. There may be gradual phonetic change in alignment. This may be likened to gradual phonetic change in segmental phonetics involving small gestural changes which over time may lead to phonological changes in the form of phonemic splits and mergers. The other kind of change in intonation may be change through pragmatic transfer of the nuclear contour from one domain onto another e.g. from one sentence/utterance type to another.

In evaluating the Alignment hypothesis, I specifically looked for phonetic alignment differences both within and between varieties. For the Transfer hypothesis however, I did not need two sentence types to differ along a parameter in this way. In fact, if statements were the 'same' as either questions or lists along one or more parameters, this could be a sign that there had indeed been transfer of question/list intonation onto a statement context. I believe it is important to look more closely at how to assess the possibility of change when two elements are the 'same'. As I mentioned in chapter 4, it is wrong to draw inferences about two elements directly from statistical analysis that does not find significant differences between them, so this is clearly a problem I would need to overcome. There is an underlying issue here that is of extreme relevance to segmental and intonational settings. Phonetic distance measures are usually quantifications of divergence between languages and varieties. They measure the number of feature values by which two varieties/languages differ. Implicit in this is the principle stemming from the Comparative method and from the Darwinian tenet of Descent with modification that only differing features between languages should be used to reconstruct language trees and the branchings in them. This approach is not restricted to studies of phonetic distance with historical orientations, though, but is widespread in purely synchronic dialectometrical studies of varieties. If two varieties share the same feature value(s), then existing methods cannot usually account for this. Hayward (2000) is one of the few to recommend quantifying convergence instead of divergence, though his proposals are general and not specifically within the field of phonetics. Thus the examination of the Transfer hypothesis in intonation highlights that segmental studies could examine the

phenomenon of convergence, including the little studied area of phonetic borrowing much further.

This approach could also be relevant in segmental studies to study semantic shifts (see section 5.2.2 above) in conjunction with segmental phonetic distances. This would link with Heggarty's method of measuring similarity in lexical semantics (McMahon et al., 2005), which takes account of the degrees of overlap in lexical semantics between languages. Thus it has the potential to show semantic shifts and meaning transfers like the following: English 'Bloom' is cognate with German 'Blume' ('flower') but English 'flower', a borrowing from French, now forms a closer meaning with the German 'Blume' (McMahon et al., 2007).

5.4.3 *Problems of using a non-linguistic framework*

A major tension among researchers of segmental distances is whether the measure should be motivated by computational simplicity or by linguistic faithfulness. Heggarty and colleagues have repeatedly criticised uses of the Edit Distance for being too crude (Heggarty, 2000b; Heggarty et al., 2005; McMahon and McMahon, 2005; Heggarty, 2006). For example, metathesis is very difficult to implement faithfully in segmental distance measures particularly those based on the Edit Distance (Kruskal, 1983), and it is sometimes avoided for this reason (Heeringa, 2004, p.125-6). The reason for this is that if the two transcriptions of an individual word from two varieties are compared, one of which has undergone metathesis, it is impossible to match up like segments simply proceeding in a left to right order. Connolly (1997) is the only one who has suggested how to apply Edit Distances to intonation, and very briefly at that. Edit Distances require discrete data but Dynamic Time Warping (DTW) is the equivalent dynamic programming procedure for continuous data (Kruskal, 1983; Kruskal and Liberman, 1983; Sankoff and Kruskal, 1983). The study of intonation leads me to advocate caution against non-linguistic approaches. In principle, it would be possible to use techniques like Dynamic Time Warping (DTW) to measure distances between intonational contours. Techniques from dynamic programming (including DTW) have been among the tools that birdsong researchers have used (Bradley and Bradley, 1983) to measure distances between different song 'dialects'. Mielke (2005) uses DTW in measuring segmental acoustic distances. However, the methods involved in dynamic programming only consider distance on a very linear level of insertions or expansions and deletions/compressions. DTW would attempt to convert one intonational contour

into another by expanding or compressing the contour along the time scale. DTW is fundamentally opposed to the notion that there are linguistically relevant points within a unit like the nuclear tone. It simply computes a score based on overall distance between two contours. This is in stark contrast to the AM treatment of the intonation contour as comprised of linguistically relevant 'events' (pitch accents, edge tones) and non-linguistically relevant 'transitions' between these events (Ladd, 2008). This aspect of DTW is extremely problematic because there is no concept of matching up like with like elements within the nuclear tone. This is also a problem with the British tradition of intonational analysis. It does not distinguish between pitch accent 'events' and boundary 'events', treating the entire nuclear tone as a unified entity. In order to match like with like, it is important to conceptualise intonational distance away from duration of the contour and syllabic content accompanying it. It is necessary on occasions to count two contours as the 'same' if one is produced on an individual word and the other stretched over several syllables (Arvaniti and Ladd, 2009). For example, I wish to consider the Belfast nuclear statement rise as the 'same' contour when produced in the 'Short' and 'Long tails' conditions respectively. This expansion is linguistically trivial but DTW would ascribe intonational distances based on durational differences alone. Therefore, it would assign distances between the Belfast 'Short' and 'Long tails' just on the basis of their having different durations in the 'tail' despite L1 and H1 points matching up in like syllables.

For the matching process in intonation, it is necessary to be able to compare a pitch accent 'event' against a pitch accent 'event' (e.g. local maxima and minima around the nuclear syllable) and a boundary 'event' against a boundary 'event' (e.g. local maxima and minima around the IP boundary). It is necessary to be able to ascribe differences/distances between two pitch accent 'events', for example, in terms of the alignment and scaling of L and/or H. Such differences represent linguistically meaningful differences, perhaps phonological differences in the choice of pitch accent or perhaps small gradient variety differences (Dalton and Ní Chasaide, 2005; Ladd et al., 2009; Mücke et al., 2009). It is not desirable to assign distances to differences between elements that are not part of pitch accent or boundary 'events' because such distances would not represent aspects of linguistic importance.

It is also unclear that a technique like DTW could give us any indications about potential processes of historical change in intonation, because it simply does not take into account any of the elements that might be relevant to change. There is

no way to decompose the DTW distance score into components indicating alignment or scaling differences. DTW's place has been in the assessment of distance between a target form and an actually produced form (e.g. in Speech recognition). This is a very different goal to my goal of deriving distances between varieties with a view to enabling a better understanding of intonational change.

The really important issue is what kind of units are involved in the first place in segments or intonation, as this will fundamentally decide which elements should be compared against which. The exploration of intonation has led me to believe that it is dubious to make intonational phonetic measurements without a concept of intonational phonology. Likewise, in segmental phonetic distances, I believe that the foundation is in segmental phonology. The consequence of this is that individual segments/phones would not be the most appropriate units of segmental measures. The Edit Distance and the **sdists** approach (chapter 2) rely on these elements, yet they may be linguistically dubious (Ladd, forthcoming). Rather, the units would be realisations of phonemes (see also Gleason, 1955, p. 158ff.). The key point is that these underlying units are not of equal size phonetically, unlike the use of segments/phones in the matching process. Simply comparing segment by segment will not always be appropriate. This is because the realisations of a given phoneme may encompass a range of phonetic qualities larger or smaller than the standard segments. I provide a couple of very simple examples from English. In varieties with post-vocalic /r/, this is realised typically as /r/ colouring [ɚ] after [ə] rather than appearing as a 'full' segment [ɹ]. In varieties without post-vocalic /r/, /r/ is realised as a null segment or by compensatory lengthening of the preceding vowel. By contrast, several English vowel phonemes are realised as diphthongs in certain varieties and thus are realised greater than a single segment. If these issues are not taken into account, there will be problems in comparing like with like.

It is evident that Heggarty (2000a) understands this. His method employs a reconstructed ancestral form for each cognate and the time slots involved allow for different combinations of segments and sub-segments. Heggarty (2000a, p.72) accepts that his time-slots "should be seen not so much as universal but as realisations of language-specific phonological distinctions, i.e. phonemes".[10] Therefore, I have come to the same conclusion as Heggarty here and I acknowledge his foresight, but I have come to this conclusion from a different perspective and

---

[10]I also included this quote from Heggarty (2000a) in chapter 2, but have included it again here because of its relevance.

this is valuable in itself. The Edit Distance and the **sdists** approach could not deal with this underlying concept of the phoneme because these approaches are rigidly tied to the single discrete segment/phone that may only be compared against another such segment (see also Kondrak, 2000). The instances in which there is not a one to one match between segment and phoneme are precisely those cases which create problems for Edit Distances and the **sdists** approaches using features that I reported on in chapter 2.

5.4.4 *Absolute vs. relative distance*

I argue that my intonational data show that relative as opposed to absolute distance is what is important. For example, the within-variety H1 alignment differences in the Belfast data were smaller than the between-variety H1 alignment difference between Belfast and Glasgow (see chapter 4 section 4.12.2 and section 5.2.4 above). Thus the relative difference has the potential to reveal where there is a phonological distinction (between Belfast and Glasgow) and where there is not (between Belfast sentence types). I disagree therefore with Laver (1994) and Kessler (2005) who believed in finding measurements of absolute distance between segments. Kessler (2005, p.250) strongly implies that ordinal distances are not "meaningful" and my data (even to some extent when converted in to ordinal distance scales in section 5.3 above) suggest the contrary. In section 5.2.4 above, I identified that it is not possible have an exact numeric conversion of segmental phonetic symbols and phonetic features. How the researcher breaks up articulatory scales of place and manner, for example, will depend to some extent on their hypotheses. Even though I had absolute acoustic measurements in intonation, combining different parameters required me to develop ordinal scales (section 5.3.1 above). Therefore, in segments and in intonation the actual distances cannot escape being somewhat relative.

5.4.5 *What is similar enough to be the same?*

I return again to the conundrum that led me on the course of intonational analysis. Dalton and Ní Chasaide (2005) queried whether the Connaught fall and the Donegal rise were instances of the same underlying form (historically and implicitly also phonologically) distinguished merely by the alignment of the peak (their "Re-alignment" hypothesis) (see chapter 3). They concluded that the Connaught and Donegal contours were not similar enough to warrant inclusion in same underlying form. They did not outline how similar the two contours would

have needed to be for this. This is treating the Connaught fall and the Donegal rise as potential allophones linking with the unresolved problem in segmental phonology of [h] and [ŋ] in English not being allophones of a single phoneme (Laver, 1994). The question though for intonation is allophones of what? It is not clear that Dalton & Ní Chasaide had defined the nature of the underlying unit.

Conceived at a purely phonological as opposed to a diachronic level, this exploration provides a way of addressing the Alignment hypothesis. If the phonemic units in intonation are the local minima and maxima associated separately with pitch accented syllables and IP-boundaries, then it would be possible to consider "Re-alignment" as affecting the $f0$ maximum within a given pitch accent unit. If H in one variety was associated to the nuclear pitch accent but in another variety was associated to the IP-boundary, then it might offer a more transparent way of rejecting this hypothesis because we would not actually be comparing like with like. This could enable us to set alignment thresholds between intonational units. Admittedly, there is still considerable disagreement over the delineation of pitch accent and boundary 'events' within AM theory, e.g. how to define 'trailing' tones of pitch accents and how to distinguish these from phrase accents or boundary tones (Grice, 1995; Grabe, 1998; Grice et al., 2000). Still, in principle, I have shown how a better understanding of the phonological units is a prerequisite to phonetic measurements. This would be relevant to segments too. Indeed the increased study of articulatory gestures (Hardcastle et al., 2010) in combination with proposals that gestures may be segmental phonological units (Browman and Goldstein, 1989, 1992) may provide headway here.

5.4.6 *What is an equivalent amount of change?*

In segmental phonetic distance measures, many researchers have not clearly considered what counts as an equivalent amount of change between two features. For example, what is an equivalent amount of height and backness change in a vowel, for example? How many places of articulation change is equivalent to a change from voiceless to voiced?

In intonation, I addressed this issue briefly by recourse to perception. I used findings for the Just Noticeable differences (JND) in alignment and scaling in proposals for how to develop discretised scales from the continuous measurements along the time and frequency scales respectively. Without recourse to perceptual

findings, it does not seem that a concept of an equivalent amount of alignment and scaling is possible.

What is an equivalent amount of change/difference in segmental phonetics? It might appear that researchers like Heggarty have addressed this issue through explicit weighting of segmental features against each other, but actually they have been addressing a different issue. Heggarty (2000a,b) weights his features by three principles. One of these is Cross-linguistic norms. The finding that most languages only make a binary phonemic distinction in voicing while a three-way contrast in place of articulation is most typical typologically is used by Heggarty (see chapter 2). This is not actually an evaluation that a change in place of articulation from labial to alveolar is a lesser change than a change in voicing. It is more about the internal organisation of both features.

Perceptual tests in segments (perhaps confusion tests/matrices) would have the potential to show how much place of articulation difference participants would consider equivalent to a voicing change or at what amount of place of articulation difference they no longer confused two segments (e.g. [f] and [θ] are known to be very confusable) (Miller and Nicely, 1955, p.347). In the segmental domain, this is problematic because the results of such tests would probably be language-specific and would also depend on position of the segment in the word among other things.[11] An acoustic measure of segmental distance could make use of Just Noticeable Differences (JNDs) (Stevens, 2002, p.228). In a diachronic setting, it would be possible to use the extensive knowledge about segmental changes, using attested stages and directions in sound changes (Labov, 1994; McMahon, 1994; Joseph and Janda, 2003). For example, [x] can become [f] without first becoming one of the intervening lingual consonants. It would be possible to document the steps involved in changes without becoming embroiled in the tricky issue of how much time a given change might take. The concept of equivalent amounts of difference along features may need to be different for consonants and for vowels. Pisoni (1973) reports that vowels are perceived much more gradiently than consonants.

This section has uncovered an issue that most researchers on segmental phonetic distances leave aside. It is much more to the forefront in intonation because of the continuous measurements along two different scales that are involved in alignment and scaling. By contrast, in articulatory phonetic distances, researchers

---

[11]JNDs in intonation are probably language-specific too.

usually measure features on discrete scales so it is not as obvious that going from step 1 to step 2 on the discrete scale for place of articulation might be intrinsically very different to going from step 1 to step 2 on the discrete scale for manner of articulation.

I acknowledge that there are problems with the use of JNDs and the other strategies mentioned. The key point is that defining equivalence is an important concern when combining different features or parameters together in an overall measure of phonetic distance. Perception and historical knowledge of segmental changes could offer ways of defining equivalence in principle.

5.4.7 Linguists meet with non-linguists

There is a clear discrepancy between the IPA/feature-based phonetic distance approaches and current research in phonetics. Recent research uses much more fine-grained articulatory techniques which concentrate much more on the level of continuous overlapping gestures e.g. Ultrasound, Electromagnetic articulography (EMA), Electromagnetic midsagittal articulometry (EMMA), Electropatalography (EPG) (Hardcastle and Laver, 1997; Hardcastle et al., 2010). Phoneticians and some other linguists increasingly view segmentation and articulatory features as somewhat dubious and there is heightened emphasis on coarticulatory phenomena (Browman and Goldstein, 1992; Port and Leary, 2005). Is there any meeting point between those interested in deriving distances and those researching phonetics for other purposes? As mentioned in chapter 1, distance measures (including phonetic distances) have increasingly attracted the attention of non-linguists. The needs of these non-linguists have little connection with current directions in phonetic theories. The most prominent need of these non-linguists is for language trees against which they may compare trees derived from archaeological, anthropological or genetic data or test hypotheses about historical movements of people (Cowlishaw and Mace, 1996; Gray and Jordan, 2000; Gray and Atkinson, 2003; Dunn et al., 2005; Gray et al., 2007; Greenhill et al., 2008; Serva and Petroni, 2008; Gray et al., 2009). Such trees can be constructed without the fine phonetic detail captured by the contemporary phonetic techniques mentioned above. Perhaps with future research, there will be increased understanding of how coarticulatory phenomena studied by these techniques link with historical change. Despite this divide between linguists and non-linguists, I do not believe there is as much of a problem here as there might appear to be. In chapter 2, I showed how existing segmental phonetic distance measures may

not really be 'phonetic' in terms of truly being able to quantify gradable articulatory or acoustic properties. In essence, the goal of trying to measure distance is completely separate from the goals of conducting phonetic measurements for other purposes. This is one reason why within-variety variability has not been addressed so far in segmental measures (section 5.4.1 above) and why measures have remained tied to the segment as a unit (section 5.4.3 above). EMMA allows researchers to get extremely detailed information the coordination of the different articulators in the vocal tract (Perkell et al., 1992). This have little direct relationship with historical questions. Therefore, if what researchers need is a broad level of phonetic/phonemic connection, then existing measures of phonetic distances can serve these general purposes. This re-iterates the point that the nature of the distance measure must relate to the researcher's hypothesis (section 5.2.4 above).

The consequence of this is that there is no intrinsic incompatibility between the direction of phonetic distances veering towards the discrete and abstract and the direction of other research in phonetics veering towards the more fine-grained. As long as the discrete, abstract approach respects Heggarty's Compatibility problem, it may serve the purpose as a valid tool for making historical judgements.

5.5 Approaching Chapter 6

This chapter has brought the segmental and intonational strands of my research on phonetic distances together. I showed the fundamental differences in approach and research questions between them while at the same time acknowledging important common concerns (particularly the Compatibility problem). I have addressed to what extent the acoustic parameters I used could be combined into an overall measure of intonational distance and I have explored the problems with trying to integrate such a measure into phylogenetic network techniques. Finally, I have turned the tables, looking afresh at segmental distances in the light of having explored intonation. This has created cohesion between two seemingly disparate issues. This is where I delimit the research I present in this thesis. Chapter 6 re-states the contribution of the thesis with reference to the segmental and intonation work I have conducted. I expose some prominent shortcomings and explain why addressing them now did not fall into the current thesis remit. The thesis concludes with suggestions for how the present work could be best extended with future research.

CHAPTER 6

Conclusions

6.1 Introduction

This thesis is about phonetic distances, what they reveal about historical change and why intonation had never before been considered in this context. Most importantly, the thesis deals with my approach to phonetic distances in intonation and my use of them to study the potential origins and trajectories of change in nuclear statement intonation in a few speakers of Belfast and Glasgow English. This chapter concludes the treatment of intonation and phonetic distances for the present and is broken down into the following main sections:

- First, I revisit the key research questions I outlined in chapter 1 and now provide answers to them (section 6.2).
- Second, I draw together the main contributions of this thesis in relation to segmental phonetic distances and intonation (section 6.3).
- Next, I discuss some of the shortcomings of this thesis and explain why addressing them did not fall within the remit of this thesis (section 6.4).
- Finally, I show how this thesis offers many original possibilities for future work either in segmental phonetic distances or in intonation (section 6.5).

6.2 Fundamental questions

In chapter 1, I laid out the following fundamental research questions of this thesis.

1. What is the difference between attempting to measure distance in segments and in intonation? How could distance in intonation be measured?
2. Can measurements of distance help to work out how intonation might have changed?
3. Is intonational change like segmental phonetic change?
4. Do intonational distance measurements allow any re-thinking of how distance is measured in segments?

6.2.1 Chapter summary

Each chapter of this thesis had a role in addressing these research questions. Chapter 1 introduced phonetic distances and intonation and showed that they had never been studied together before. Chapter 2 gave extensive treatment of how phonetic distances have been measured in segments. Drawing on experimental comparisons of a few existing measures, I demonstrated some problems with these and highlighted which elements of segmental measures would be relevant to measuring distances in intonation. In Chapter 3, I explored intonational change theoretically, taking nuclear statement intonation in Belfast and Glasgow English as a central case study. Chapter 4 applied phonetic measurements of intonational alignment and scaling to contemporary Belfast and Glasgow English data. I used these measurements to evaluate two hypotheses (the Alignment and Transfer hypotheses) on the origin of nuclear statement intonation in a few speakers of these varieties. Chapter 5 discussed some key similarities and differences in the underlying goals of segmental and intonational distance measurements, showed how the parameters of alignment and scaling could be combined into an overall intonational distance measure, and suggested how intonational measurements may shed new light on approaches to segmental distances.

6.2.2 Responses to fundamental questions

The main differences between measuring distance in segments and in intonation are:

- Intonational measurements involve continuous acoustic measurements whereas the majority of existing segmental measures use discrete articulatory feature values.

- Intonational measurements are applied to several utterances in different conditions within each variety as well as between varieties. Segmental approaches tend to make cross-variety comparisons only, using isolated transcribed tokens of individual words.

In chapter 5, I outlined ways in which segmental acoustic measurements and within-variety variability could be used to study segmental change. Of course, there are several other differences. For example, researchers tend to use segmental measures in an historical context to confirm ancestral relationships between languages and varieties rather than to chart the trajectory of phonetic sound change. In intonation, it was precisely the study of the path of change that interested me in relation to Belfast and Glasgow English.

My approach to measuring distance in intonation took the nuclear tone as a unit. This is unlike the cognate (used in conjunction with segmental measures) in that it does not have historical status. The nuclear tone is a well-defined unit, long-established as being the most important part of the intonation contour and it enabled me to match like with like elements within it. Matching like with like is a central problem for segmental and intonational measurements. Within this unit, I isolated tonal targets of an $f0$ minimum (L1) and $f0$ maximum (H1) to be compared along the acoustic parameters of alignment and scaling. These parameters are central, but not exclusive to AM theory. I used intonational measurements along these parameters to evaluate two hypotheses, the Alignment and Transfer hypotheses. These intonational measurements allowed me to isolate crucial differences between small datasets of Glasgow and Belfast English, suggesting that they have undergone separate trajectories of change. Although I explored an overall intonational distance measure in chapter 5, individual measurements of alignment and scaling were sufficient for me to argue that the Glasgow data supported the Alignment hypothesis and the Belfast data, the Transfer hypothesis. This links with a central argument from chapter 2, that isolating the segmental features most relevant to one's hypothesis may be more appropriate than combining several features together.

The experimental results presented in chapter 4 showed that intonational change could happen in a manner similar to segmental change but also that it could happen in a very different way too. The Alignment hypothesis predicted that intonational change would be phonetically gradual with subtle articulatory, acoustic

and perceptual shifts eventually giving way to perceptual and phonological re-organisation (see chapter 3).

The Transfer hypothesis predicted that intonational change is not like segmental phonetic change. Rather it proposed that intonational change would involve pragmatic/discourse transfer of the nuclear intonation contour of one domain to another. The domain instigating the transfer could be question intonation or continuation/list intonation.

Extending phonetic distance measurements into intonation has given valuable new perspectives on segmental phonetic distance measurements. Within-variety synchronic variability in the Glasgow data was essential to explaining potential diachronic change. Segmental methods have great scope to gain insights about the progress of sound change by moving away from individual isolated word tokens per variety. Given that phonetic measurements can suggest types of change other than gradual phonetic change, segmental researchers may be encouraged to explore other kinds in segmental phonetic change with segmental distances too. My approach to intonational measurements has shown there are problems with using a non-linguistic framework (often favoured by segmental researchers) and such a framework cannot incorporate concepts essential to matching up like with like.

6.3 Contributions

In chapter 1, I stated the central argument of this thesis, that intonation deserved a place within the context of phonetic distance which had thus far only been studied in segments. I argue that the subsequent chapters have confirmed intonation's place in this context.

The most fundamental contribution of this thesis is that intonational measurements have enabled me to argue that nuclear statement intonation contours in a few speakers of Belfast and Glasgow English may have different origins.

This tackles the key under-explored research topics, outlined in chapter 1, that I aimed to address in this thesis.

- the treatment of *intonation* in the context of measurements of phonetic distance;
- mechanisms of sound change in intonation;

- an explicit comparison of intonation between Belfast and Glasgow English;
- how treatments of intonational distance may add to the understanding of segmental distance.

First, the Belfast and Glasgow English findings show that phonetic measurements of contemporary synchronic data could indeed help to answer historical questions in intonation as well as in segmental data. Given the paucity of actual historical evidence in intonation, this approach may be one vital way of accessing information about historical change in intonation. Second, these intonational measurements have also pointed to two types of sound change in intonation, gradual phonetic alignment change and pragmatic/discourse transfer. Thirdly, the context of phonetic distances offered a platform for the first direct comparison of Belfast and Glasgow English intonation. Links between these two varieties are long-standing, through reports that they share a distinctive nuclear rising phenomenon on statements (Cruttenden, 1995, 1997; Ladd, 2008). The intonational measurements I conducted showed that whatever the impressionistic similarities between the intonation of these varieties may be, the few speakers in the present study from Belfast and Glasgow English respectively did not display a unified phenomenon.

There are a number of other contributions of this thesis. The Glasgow English data showed support for the Alignment hypothesis. Within-variety allophony in the alignment of H1 between the 'Short' and 'Long tails' conditions was the primary source of support. In the 'Short tails' H1 was aligned within the nuclear syllable and in the 'Long tails' H1 was aligned in a postnuclear syllable. There was a significant difference in the alignment of H1 according to 'Tail'. The synchronic alignment shift may indicate a process of diachronic change whereby H1 was aligned in the nuclear syllable originally in 'Short' and 'Long tails'. However, it has drifted to the right along with the increase in the number of following unstressed syllables. This more rightward alignment has become entrenched in the Glasgow 'Long tails' data. Combined with this was the presence of a final low in the Glasgow data but not in the Belfast data.

The Belfast English data do not show support for the Alignment hypothesis because there is no difference in the alignment of H1 between the 'Short' and 'Long tails'. The Belfast data show more support for the Transfer hypothesis. Belfast statements are closer in the alignment of H1 and in scaling of H1 to Belfast lists

than to Belfast questions, suggesting that the transfer may come from continu-ation/list intonation. This is an extended perspective on intonational transfer. Previously, researchers had only considered the possibility that transfer could come from question intonation (Bolinger, 1978). Nevertheless, the H1 scaling differences between statements and questions may be paralinguistic and thus not as central in decisions about the nature of intonational change as alignment differences.

Before embarking on intonational measurements, I needed to understand in de-tail the mechanisms of existing segmental distance measures (Chapter 2). Exper-imental work I conducted to address this revealed that:

1. Phonetic detail in the form of differing feature values may fail to make an impact on overall distance scores between varieties.
2. A measure composed of only the features relevant to one's hypotheses may be more appropriate than a measure including as many features as possible.
3. Non-linguistic aspects of the measuring process may actually erase attempts to reflect gradient phonetic differences between varieties.

The second of these points turned out to be very relevant to intonation. In eval-uating the Alignment and Transfer hypotheses, I also found the need to focus on the parameters directly relevant to these hypotheses.[1] The third point also had connections with intonational measurements, as non-linguistic frameworks used with intonation could give distances that would be very dubious (chapter 5). By including both segmental and intonational distance measurements in this thesis, I have been able to emphasise their differences while at the same time showing that they have valuable contributions to make to each other.

In chapter 5, I explored how the intonational measurements of alignment and scaling could be combined into an overall measure of intonational distance. I also provided a representation of the resulting distances using NeighborNet, a phylogenetic network algorithm increasingly used with segmental phonetic dis-tance measurements (Maguire et al., 2010). I discussed problems with weight-ing intonational parameters for this overall score. Although the measurements I derived complemented my arguments in chapter 4 about the Alignment and Transfer hypotheses to some extent, such combined measurements were not ac-tually necessary for the evaluation of these hypotheses. Still, this represents one

---

[1]In any case, I only used a few parameters (alignment and scaling of L1 and H1) in my into-national measurements to begin with.

of the first attempts at an overall measure of intonational distance and the first incorporation of intonation in the NeighborNet framework.

6.3.1 *What is it for?*

I see this thesis as particularly relevant to a number of fields. First among them would be the study of historical change in intonation. Separately, the attempt to consider phonetic distance in intonation would provide a point of reference to those wishing to derive intonational distances for other purposes (including clinical purposes). The study of Belfast and Glasgow English intonation adds to the growing body of work on phonetic alignment and scaling of closely related varieties (Arvaniti and Garding, 2007; Ladd et al., 2009; Mücke et al., 2009). It also adds to the study of varieties more generally. In chapter 2, I discussed how studies of dialectometry tend to remain restricted to comparisons of isolated transcriptions of individual words. Acoustic intonational data of several utterances per variety in the case of Belfast and Glasgow clearly represent a move away from this. This thesis also suggests that intonation may be a new application for a combined measure of phonetic distance for incorporation with phylogenetic network methods like NeighborNet.

6.4 Shortcomings

Naturally, this thesis contains shortcomings. I now expose some of the most prominent ones. I also explain why addressing them was not necessary for the crucial arguments of the present thesis. The segmental experimental work I presented in chapter 2 could have been much expanded to provide a much more comprehensive comparison of different measures of phonetic distance over a broad set of varieties. Instead, it remains a small-scale study of just four different measures, ten varieties, and thirty words. The **sdists** framework which I used even for the feature methods was restrictive. I did not include the method of Heggarty (2000a) for comparison with the other measures, despite Heggarty's Compatibility and Quantification problems being very influential on my later attempts to measure intonational distance. Nor did I include acoustic segmental distance measures in this small-scale study. Yet the underlying goal of this thesis was not to expand the study of segmental phonetic distance or to provide an in-depth comparison of existing measures. The goal of this thesis was to move away from these segmental measurements and to begin tackling phonetic distance measurements in intonation.

The segmental comparisons in chapter 2 did not lead to any comparison of different phylogenetic methods. I only used NeighborNet. The goal of this thesis was not similar to that of Nakhleh, Warnow, Ringe and Evans (2005), who evaluated the performance of different tree-building algorithms. By contrast, I simply raised the question whether intonational measurements could have a place within these tree and network methods, not which methods might be better than others.

Though this thesis argues that Belfast and Glasgow nuclear statement intonation contours may have separate origins, in fact the Glasgow data shows synchronic allophonic variation between 'Short' and 'Long tails'. I used this to make historical inferences. More generally, it is problematic that at present I am unable to back up the key synchronic findings from chapter 4 with evidence from archive recordings or Apparent Time data. The aim of the thesis though was to see what potential acoustic measurements of contemporary data had in evaluating historically-oriented hypotheses in intonation, not to scour all possible sources of archive data nor to conduct quantitative sociolinguistic work (the domain of Apparent Time work). Indeed Kroch (2001, p.15) separately argues that vernacular data from the past

> show how delicate the interpretation of diachronic evidence is and suggest that it is much easier to explain the past through the study of the present than vice versa.

The main experiment in chapter 4 (Experiment II) involved just a few speakers each of Belfast and Glasgow English and all were university students. There were problems with eliciting enough appropriate data from Glasgow, leaving me to rely on one primary speaker from this variety. Using the HCRC Map Task corpus, I concluded that this speaker was indeed representative of a much wider body of Glasgow English speakers. A better sociolinguistic understanding of Glasgow English would no doubt have helped me to collect more usable Glasgow data, but a sociolinguistic study of any of the English varieties to which I refer in this thesis was not one of my aims. Despite larger-scale intentions, I only included Edinburgh English in a brief comparison with Glasgow wh question data. The study of intonational change has thus focused on a very small number of varieties and just the single context of the phenomenon of nuclear statement

rising intonation. The similar phenomenon in Donegal Irish and tentative exploration of change therein (Dalton and Ní Chasaide, 2005; Dalton, 2007) was an important motivation for me in examining potential change in Belfast and Glasgow English. However, I did not make any direct comparisons with the Irish data. I also did not make large-scale use of the extensive IViE corpus (Grabe et al., 2001). In chapter 4, I made it clear that I needed specially designed materials, specifically the need for a contrast between 'Short' and 'Long tails' sets. Neither the Irish language data nor the IViE data contain a systematic set of utterances with enough unstressed syllables in the 'tail' for appropriate comparison with my 'Long tails' set.

I acknowledge that there are some methodological concerns with using the $f0$ minimum and maximum instead of turning points ('elbows') to represent the beginning and end of the nuclear rise. In chapter 4, I showed problems too with current methods for locating 'elbows' and this thesis did not have the goal of improving on existing algorithms for this purpose. I also did not have a specific measure of the final low in the Glasgow data and there is little consensus on how scaling should be normalised (Ladd, 2008). Despite these issues, I was able to show robust differences between Belfast and Glasgow in the scaling of the IP-final vowel and within-variety differences in H1 scaling between statements and questions. Therefore, the normalised scaling procedure I used was sufficient to allow me to assess key aspects of the Alignment and Transfer hypotheses respectively. Alignment and scaling (along with $f0$ excursion) were the only parameters I used in the intonational phonetic measurements. There were several other parameters I could have used including the non-$f0$ parameters which I briefly explored at the end of chapter 4. One of the central arguments of this thesis though is for the use of the features/parameters directly relevant to one's hypotheses. I believe that I have amply demonstrated (chapter 4) that alignment and scaling are of central relevance to the Alignment and Transfer hypotheses.

The overall measure of intonational distance presented in chapter 5 is a slender outline and would need several refinements, especially in developing an approach to weighting different parameters. Likewise, my suggestions for how the exploration of intonation could aid approaches to measuring segmental distance would need to be much more fully worked through to be ready for implementation. Yet this thesis represents the start of integrating intonation into a phonetic distance framework and offers scope for these refinements to follow in future work.

6.5   Future directions

Many of the contributions and shortcomings of this thesis could consolidate the central argument of this thesis, that intonation deserves a place in a phonetic distance framework. A future study could attempt to replicate the finding of the crucial difference between Glasgow and Belfast intonation on a much larger-scale and could refine the non-$f0$ parameters briefly introduced at the end of chapter 4. This new study would address the sociolinguistic situation in Glasgow as a central concern, through a large-scale treatment of social class, gender and speech style differences. Such a study could also incorporate the other varieties of English included in the 'Urban North British' (UNB) group of varieties and Irish to see whether they more closely resembled Glasgow or Belfast. Descriptions of Liverpool English intonation from Knowles (1984) suggest a closer connection with the Glasgow pattern I found in my data. By contrast, Donegal Irish appears to behave much more like Belfast English (Dalton and Ní Chasaide, 2005; Dalton, 2007). It would be extremely valuable to compare High Rising Terminals (HRTs) from Antipodean or U.S. English directly against UNB data, with underlying questions about the origin and development of HRTs. In this context, a larger analysis of Glasgow English Map task data would be valuable to see if there really are the two separate patterns, the HRT-style nuclear contours with no final low and the more typical Glasgow nuclear contours with a final low (see chapter 4).

The inclusion of these other varieties would create a much greater set of varieties from which to make pairwise comparisons using an overall measure of intonational distance. Refinements to the overall intonational measure I outlined in chapter 5 and to existing segmental measures (see chapter 2) would enable direct comparison of intonational and segmental distances on the same data. It would then be possible to compare the resulting connections between varieties from segmental and intonational distances respectively, using NeighborNet. This avenue would enable the researcher to test on a greater scale whether individual parameters targeting a hypothesis are indeed more appropriate for intonation than a combined measure of distance.

A categorical perception test would be necessary to examine the phonological status of the H1 scaling differences between statements and questions. Such a test would also enable me to assess the implications of categorical contrasts and

paralinguistic variability on intonational change more thoroughly. The possibility of transfer of continuation intonation onto a statement context could be much further studied by the examination of turn-holding and yielding devices in Map task data and topic-medial and final intonation in spontaneous spoken discourse of an individual speaker. This would be a strong addition to the list sentences in the Sentence reading task of Experiment II. It would be important to examine the Alignment and Transfer hypotheses in other languages than English or Irish. Particularly appropriate would be the Norwegian data examined by Hognestad (2006). Hognestad argues for alignment change leading to extremely delayed peaks in some Norwegian varieties, especially Haugesund. The location of these extremely delayed peaks may be similar to the location of H1 in the Belfast English data, in which I have argued against against alignment change and rather for the Transfer hypothesis.

Probing intonational change with these two hypotheses (or other ones more suitable for the data) would be important in other contexts too. Colantoni and Gurlekian (2004) posit a leftwards alignment shift diachronically in Buenos Aires Spanish. Elordieta and Calleja (2005) and Hualde (2007) also argue for a leftwards alignment shift in certain Basque varieties of Spanish and in Basque language varieties themselves respectively. These studies all invoke contact as a motivation for intonational change. The notion of contact as transfer would be a valuable direction in which to take the Transfer hypothesis.

Returning to Belfast and Glasgow English, archive recordings of Northern Irish English and Irish are available dating from the early to mid 20th century at the National Folklore Collection, University College Dublin and at the Ulster Folk and Transport Museum, Hollywood, Co. Down (Tape-recorded survey of Hiberno-English (Adams et al., 1985)). An in-depth exploration of these could provide valuable supplementary material for evaluating hypothesis about intonational change and indeed perhaps for tracking the progress of change. A carefully planned Apparent Time study could also be very fruitful. Contemporary techniques of measuring intonation together with articulation (e.g. Mücke et al., 2009) as well as deriving perceptual intonational distances (see Gussenhoven and Rietveld, 1991) would offer further ways to extend this treatment of intonational distance and change in Belfast and Glasgow English.

6.6 Conclusion

This chapter has reiterated the fundamental research questions underpinning the entire thesis and has drawn strands from previous chapters to provide answers to these questions (section 6.2). I have shown what the thesis has added to knowledge about intonation in Belfast and Glasgow English and about the relationship between phonetic distances and sound change in segments and in intonation (section 6.3). Shortcomings are unfortunate but inevitable (section 6.4) but they can be overcome with future work (section 6.5).

> Speaking English means doing two things at once: saying the words, and saying the melody (Silverman and Pierrehumbert, 1990, p.72).[2]

Words versus 'melody' was the fundamental distinction with which this thesis began. Specifically, I intended to move away from measuring phonetic distance between vowel and consonant segments in isolated transcribed words to measuring distance in 'melody' or rather, intonation. I have had the goal of linking intonational measurements to hypotheses about intonational change. I have done this in the special context of nuclear statement intonation in Belfast and Glasgow English. If I told you now that nuclear statement intonation in the few speakers that I studied of Belfast and Glasgow English respectively is different and reflects different processes of intonational change, would you believe me? Your answer decides if this thesis has been convincing or not.

---

[2]A reminder of the quote at the opening of chapter 1 of this thesis

Appendix A

Glossary

Segment A discrete representation of an individual vowel or consonant. Contrast with Suprasegmental phenomena, which encompass larger stretches of speech than individual vowels and consonants. Intonation has typically been considered as suprasegmental.

Lexicostatistics A technique for measuring the degree of lexical distance between languages and varieties (Swadesh, 1972). It uses lists of 100/200 basic meanings and counts the number of meanings for which a given pair of languages/varieties have cognate words.

Edit Distance An algorithm which converts one computational string into another using the least number of insertions, deletions and substitutions of discrete elements (Wagner and Fischer, 1974). This algorithm is often used to compute distances between phonetic transcriptions, where it involves insertions, deletions and substitutions of phonetic symbols (Heeringa, 2004). It is also known as the Levenshtein distance.

Phylogenetic techniques Tree- and Network-drawing procedures designed to study evolutionary history in biology. They have been increasingly used in Historical linguistics to study ancestral connections between languages, borrowing and phonetic distances (McMahon and McMahon, 2005).

Intonation The use of pitch (or its acoustic correlate, fundamental frequency ($f0$)) along with elements of loudness (intensity) and duration to make utterance-level meaning contrasts (Ladd, 2008).

Lexical tone The use of pitch (or $f0$) to make lexical contrasts. For example, in Wu Chinese, /di/ means 'lift' when spoken with Tone I (low level), 'younger brother' with Tone II (low rising) and 'field' with Tone III (low falling) (Chen, 2000, p. 6-8).

Lexical accent A more restricted lexical use of pitch e.g. in Stockholm Swedish (Bruce, 1977). In general, languages with lexical accent not only have a smaller inventory of accents than languages with lexical tone but also use accent on a much smaller set of syllables than in languages with lexical tone.

Pitch accent Part of the intonation contour associated with particular stressed syllables marking these syllables as prominent. Pitch accents usually contain a local minimum or maximum in the $f0$ contour (Ladd, 2008, p.49)

Intonational Phrase (IP) When speaking continuously, speakers break their speech down into chunks, delimited by pauses, by segmental lengthening of vowels and by specific pitch patterns (de Pijper and Sanderman, 1994; Shattuck-Hufnagel and Turk, 1996). These chunks have been argued to correspond to prosodic units and there is a hierarchy of these units. The Intonational Phrase (IP) is one of the largest units in this hierarchy, equivalent to or just below the Utterance. Complete intonation contours work within this unit (Ladd, 2008, ch.8). The IP often corresponds to a simple sentence with a single clause. However, IPs are not the same thing as syntactic phrases and do not always correspond neatly with syntactic phrases.

Nuclear tone The most prominent stressed syllable and subsequent syllables extending until the end of the IP have long held special status in intonational analysis. This is particularly so in the British tradition of intonational analysis (Crystal, 1969; Cruttenden, 1997). There the most prominent stressed syllable is called the 'Nucleus' or 'Nuclear syllable'. The following syllables are grouped together to form the 'Tail'. The entire unit encompassing the nucleus and the 'tail' is the nuclear tone.

Linear Mixed-effect models Statistical techniques enabling the researcher to model data using fixed and random effects (Baayen, 2008; Baayen et al., 2008). Fixed effects are central to the researcher's hypothesis and could be measured again at the same levels in future studies, e.g. Sentence type in Experiment II of this thesis (chapter 4). Random effects are not expected to affect the data systematically and could not be measured again at the same levels in future studies, e.g. speaker/participant. Linear Mixed-effect models are more robust to violations of the assumptions of more traditional procedures like ANOVA. Significance may be assessed using Markov Chain Monte Carlo (MCMC) sampling.

APPENDIX B

This appendix contains material relevant to chapter 2.

Feature	Value	Meaning
vowel	0	no
	1	yes
advancement	1	front
	2	central
	3	back
height	1	close
	2	near-close
	3	close-mid
	4	central
	5	open-mid
	6	near-open
	7	open
roundedness	0	no
	1	yes
long	0	short
	0.5	half-long
	1	long

Table B.1: Table of phonetic features and their numerical values for the 'Heeringa version' (vowels) (Heeringa, 2004, p.44)

Feature	Value	Meaning
consonant	0	no
	1	yes
place	1	bilabial
	2	labiodental
	3	dental
	4	alveolar
	5	postalveolar
	6	retroflex
	7	palatal
	8	velar
	9	uvular
	10	pharyngeal
	11	glottal
manner	1	plosive
	2	nasal
	3	trill
	4	tap or flap
	5	fricative
	6	lateral fricative
	7	approximant
	8	lateral approximant
voice	0	no
	1	yes

Table B.2: Table of phonetic features and their numerical values for the 'Heeringa version' (consonants) (Heeringa, 2004, p.44)

Name of feature	Feature value	Interpretation of feature value
Glottalic	1	Ejective
	0.5	Pulmonic
	0	Implosive
Velaric	1	Click
	0	non-Click
Voice	1	Glottal stop
	0.8	Laryngealized
	0.6	Voice
	0.2	Murmur
	0	Voiceless
Aspiration	1	Aspirated
	0.5	Unaspirated
	0	Voiced
Place	1	Labial
	0.9	Dental
	0.85	Alveolar
	0.8	Post-alveolar
	0.75	Pre-palatal
	0.7	Palatal
	0.6	Velar
	0.5	Uvualar
	0.3	Pharyngeal
	0	Glottal
Constrictor	1	Labially constricted
	0.9	Dentally constricted
	0.85	Apically constricted
	0.75	Laminally constricted
	0.6	Dorsally constricted
	0.3	Radically constricted
	0	Glottally constricted
Stop	1	Stop
	0.9	Fricative
	0	Approximant
Nasal	1	Nasal(ized)
	0	Non-nasal(ized)
Lateral	1	Lateral
	0	Non-lateral
Trill	1	Trill
	0	Non-trill
Tap	1	Tap
	0	Non-tap
Sulcal	1	Narrow groove
	0.8	Broad groove
	0	No groove

Table B.3: Table of phonetic features and their numerical values for the Connolly method (part I) (Connolly, 1997, p.282-3)

Height	1	High
	0.75	High-mid
	0.5	Mid
	0.25	Low-mid
	0	Low
Back	1	Back
	0.5	Central
	0	Front
Round	1	Rounded
	0	Unrounded
Wide	1	Wide
	0.5	Normal
	0	Narrow
Retroflexed	1	Retroflexed
	0	Non-retroflexed
Length	1	Of long vowel
	0.6	Of shortened long vowel
	0.5	Of short vowel or long consonant
	0.4	Of typical consonant
Double	1	Double articulation
	0	Single articulation

Table B.4: Table of phonetic features and their numerical values for the Connolly method (part 2)

Almeida and Braun (1986) method (my translation)

Vowels:

- Two points for a single difference each along two dimensions;
- Differences in the same dimension of two or more symbols score 2 points;
- Secondary articulation (nasality): 1 point;
- Length: long vs. short vowel: 1 point, short vs. half-long 0.5 point;
- Voice quality difference: 1 point;
- Maximal 8 points (2 tongue height + 2 backness + 1 round + 1 length + 1 nasality + 1 voice quality);
- This is scaled back to a maximum of 3 points (somewhat like a log scale).
- Monophthong vs. diphthong: 3 points;
- Vowel vs. 'silence': 3 points.

Consonants:

- 1 point for neighbouring differences along place, manner or voicing;
- As with vowels in each dimension, a difference of two points or more: 2 points only e.g. [ɸ] and [θ];
- Difference in two dimensions: 2 points e.g. [p] vs. [β];
- Differences in more than two dimensions: 3 points;
- Diacritics: 0.5 point;
- Aspiration: 0.5 point;
- Missing segments: 3 segments;
- Secondary articulation and /r/ colouring: 1 point;
- Ceiling overall score of 3 points.

Table B.5: Numerical values for the phonetic features in the Almeida & Braun method

- R function used to calculate the basic Edit Distance:

$$sdists(x, weight = c(1, 0, 1)) \tag{B.1}$$

- R function used to calculate the distances in each of the feature methods:

$$sdists(x, method = \text{``aw''}, weight = y) \tag{B.2}$$

y is a .csv file (converted to a matrix in R) containing distances between individual segments for each feature method respectively.

Table B.6: R functions used to calculate distances

brother	cold	daughter	eye	foot	heart
holy	home	horn	ice	long	mother
mouth	north	over	right	storm	swear
white	word	one	two	three	four
five	six	seven	eight	nine	ten

Table B.7: List of the 30 words used in the small-scale study, a subset of the 110 Germanic cognates in McMahon et al. (2005-07)

Variety pairs	mean distance	% distance
Ge vs. Am	0.693 (s.d. 0.235)	69
Ge vs. RP	0.662 (s.d. 0.249)	66
Ge vs. Sc	0.722 (s.d. 0.217)	72
Ge vs. Bu	0.7 (s.d. 0.227)	70
Am vs. RP	0.215 (s.d. 0.182)	22
Am vs. Sc	0.381 (s.d. 0.221)	38
Am vs. Bu	0.534 (s.d. 0.266)	53
RP vs. Sc	0.38 (s.d. 0.224)	38
RP vs. Bu	0.551 (s.d. 0.235)	55
Sc vs. Bu	0.469 (s.d. 0.244)	47

Table B.8: Means, sds and % distances in the Basic Edit Distance for five Germanic varieties (see chapter 2 for explanation of the abbreviations)

For the calculation of % distance scores in the feature methods, I needed to create a scale between 0 and 1. I chose the highest scoring word distance between a pair of varieties to represent the top end of the scale. I divided all the mean distances by this value, then multiplied by 100 to get a % score. For the 'Heeringa version', the highest score was for the word 'two' and was 10.67. For the Connolly method, the highest score was for the word 'four' and was 3.33. For the Almeida & Braun method, the highest score was scored equally between 'four' and 'over' and was 3.75.

Table B.9: Calculations of % distance scores in the feature methods

Variety pairs	mean distance	% distance
Ge vs. Am	3.316 (s.d. 1.977)	31
Ge vs. RP	2.905 (s.d. 1.826)	27
Ge vs. Sc	3.939 (s.d. 1.99)	37
Ge vs. Bu	3.366 (s.d. 1.788)	32
Am vs. RP	1.017 (s.d. 1.523)	10
Am vs. Sc	1.466 (s.d. 1.33)	14
Am vs. Bu	2.981 (s.d. 2.579)	28
RP vs. Sc	1.991 (s.d. 1.572)	19
RP vs. Bu	3.22 (s.d. 2.549)	30
Sc vs. Bu	2.468 (s.d. 2.632)	23

Table B.10: Means, sds and % distances in the 'Heeringa version' for five Germanic varieties

Variety pairs	mean distance	% distance
Ge vs. Am	1.131 (s.d. 0.661)	34
Ge vs. RP	0.97 (s.d. 0.631)	29
Ge vs. Sc	1.29 (s.d. 0.598)	39
Ge vs. Bu	1.188 (s.d. 0.644)	36
Am vs. RP	0.321 (s.d. 0.431)	10
Am vs. Sc	0.484 (s.d. 0.414)	15
Am vs. Bu	0.948 (s.d. 0.67)	28
RP vs. Sc	0.655 (s.d. 0.427)	20
RP vs. Bu	0.985 (s.d. 0.643)	30
Sc vs. Bu	0.832 (s.d. 0.659)	25

Table B.11: Means, sds and % distances in the Connolly method for five Germanic varieties

Variety pairs	mean distance	% distance
Ge vs. Am	1.607 (s.d. 0.764)	43
Ge vs. RP	1.497 (s.d. 0.748)	40
Ge vs. Sc	1.787 (s.d. 0.797)	48
Ge vs. Bu	1.543 (s.d. 0.639)	41
Am vs. RP	0.463 (s.d. 0.51)	12
Am vs. Sc	0.837 (s.d. 0.521)	22
Am vs. Bu	1.424 (s.d. 0.866)	38
RP vs. Sc	0.917 (s.d. 0.49)	24
RP vs. Bu	1.44 (s.d. 0.795)	38
Sc vs. Bu	1.309 (s.d. 0.921)	35

Table B.12: Means, sds and % distances in the Almeida & Braun method for five Germanic varieties

Variety pairs	% distance
Ge vs. Am	42
Ge vs. RP	40
Ge vs. Sc	40
Ge vs. Bu	36
Ge vs. No	38
Ge vs. Sw	46
Ge vs. Li	41
Ge vs. Ty	40
Ge vs. Be	40
Am vs. RP	6
Am vs. Sc	8
Am vs. Bu	20
Am vs. No	43
Am vs. Sw	46
Am vs. Li	12
Am vs. Ty	10
Am vs. Be	5
RP vs. Sc	12
RP vs. Bu	22
RP vs. No	45
RP vs. Sw	48
RP vs. Li	7
RP vs. Ty	5
RP vs. Be	8
Sc vs. Bu	18
Sc vs. No	43
Sc vs. Sw	43
Sc vs. Li	17
Sc vs. Ty	12
Sc vs. Be	5
Bu vs. No	40
Bu vs. Sw	42
Bu vs. Li	27
Bu vs. Ty	24
Bu vs. Be	18
No vs. Sw	20
No vs. Li	51
No vs. Ty	45
No vs. Be	43
Sw vs. Li	54
Sw vs. Ty	46
Sw vs. Be	45
Li vs. Ty	10
Li vs. Be	14
Ty vs. Be	10

Table B.13: % distances in the 'Heeringa version' for the numerals 1-10 for 10 Germanic varieties. See chapter 2 for explanation of the abbreviations.

Variety pairs	% distance
Ge vs. Am	50
Ge vs. RP	45
Ge vs. Sc	45
Ge vs. Bu	43
Ge vs. No	41
Ge vs. Sw	53
Ge vs. Li	44
Ge vs. Ty	50
Ge vs. Be	43
Am vs. RP	5
Am vs. Sc	13
Am vs. Bu	24
Am vs. No	52
Am vs. Sw	54
Am vs. Li	8
Am vs. Ty	15
Am vs. Be	8
RP vs. Sc	14
RP vs. Bu	24
RP vs. No	53
RP vs. Sw	54
RP vs. Li	4
RP vs. Ty	10
RP vs. Be	9
Sc vs. Bu	24
Sc vs. No	51
Sc vs. Sw	52
Sc vs. Li	16
Sc vs. Ty	23
Sc vs. Be	11
Bu vs. No	44
Bu vs. Sw	50
Bu vs. Li	26
Bu vs. Ty	33
Bu vs. Be	18
No vs. Sw	26
No vs. Li	56
No vs. Ty	57
No vs. Be	48
Sw vs. Li	57
Sw vs. Ty	54
Sw vs. Be	50
Li vs. Ty	13
Li vs. Be	12
Ty vs. Be	18

Table B.14: % distances in the Connolly system for the numerals 1-10 for 10 Germanic varieties

Variety pairs	% distance
Ge vs. Am	63
Ge vs. RP	60
Ge vs. Sc	55
Ge vs. Bu	53
Ge vs. No	47
Ge vs. Sw	61
Ge vs. Li	62
Ge vs. Ty	61
Ge vs. Be	55
Am vs. RP	7
Am vs. Sc	20
Am vs. Bu	35
Am vs. No	58
Am vs. Sw	61
Am vs. Li	11
Am vs. Ty	21
Am vs. Be	14
RP vs. Sc	20
RP vs. Bu	35
RP vs. No	58
RP vs. Sw	62
RP vs. Li	7
RP vs. Ty	16
RP vs. Be	16
Sc vs. Bu	32
Sc vs. No	59
Sc vs. Sw	62
Sc vs. Li	24
Sc vs. Ty	29
Sc vs. Be	16
Bu vs. No	54
Bu vs. Sw	60
Bu vs. Li	37
Bu vs. Ty	43
Bu vs. Be	27
No vs. Sw	28
No vs. Li	62
No vs. Ty	63
No vs. Be	58
Sw vs. Li	65
Sw vs. Ty	62
Sw vs. Be	62
Li vs. Ty	18
Li vs. Be	20
Ty vs. Be	25

Table B.15: % distances in the Almeida & Braun system for the numerals 1-10 for 10 Germanic varieties

Feature systems	Segment type	Correlation coefficient	p value
Heeringa vs. Almeida & Braun	vowel	0.791	$p < 0.002$
Heeringa vs. Connolly	vowel	0.727	$p < 0.002$
Almeida & Braun vs. Connolly	vowel	0.817	$p < 0.002$
Heeringa vs. Almeida & Braun	consonant	0.51	$p < 0.002$
Heeringa vs. Connolly	consonant	0.507	$p < 0.002$
Almeida & Braun vs. Connolly	consonant	0.477	$p < 0.002$

Table B.16: Mantel tests on the vowel and consonant parts of the feature methods. Mantel tests were conducted using the **ncf** package in R with 10000 replications. The p value is Bonferroni corrected.

APPENDIX C

Appendix C also contains material relevant to chapter 2.

H1.0
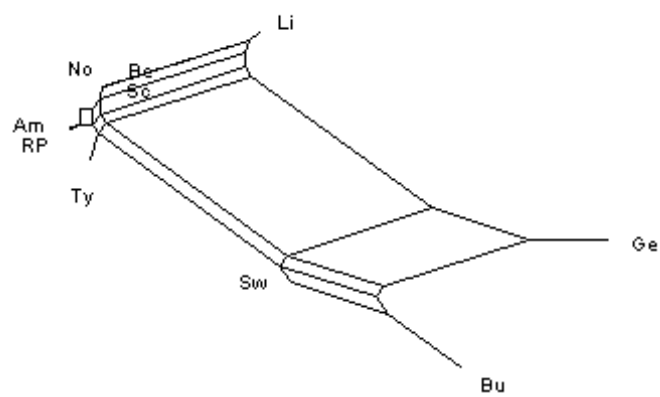
Figure C.1: NeighborNet diagram for the 'Heeringa version' for the numeral 'two'. See chapter 2 for explanation of the abbreviations.
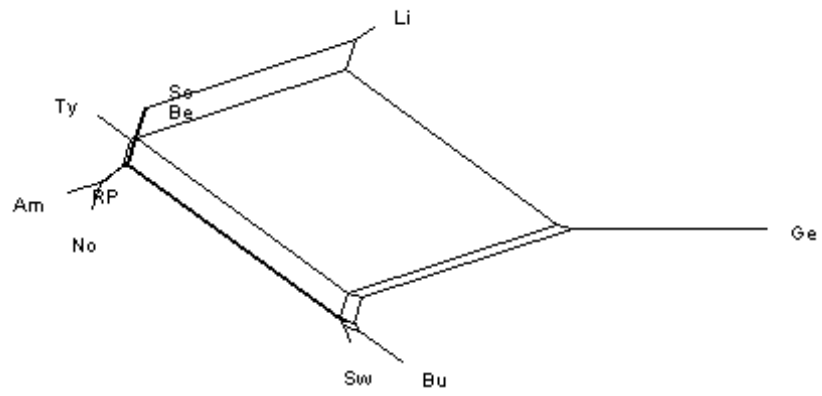
Figure C.2: NeighborNet diagram for the Connolly method for the numeral 'two'
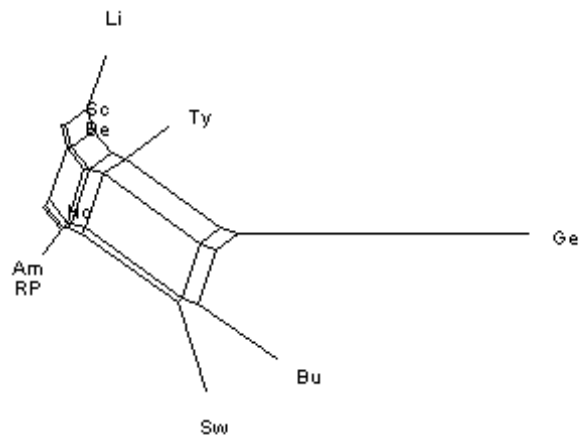
Figure C.3: NeighborNet diagram for the Almeida & Braun method for the numeral 'two'

H1.0
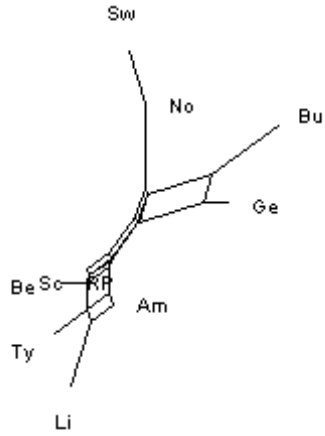
Figure C.4: NeighborNet diagram for the 'Heeringa version' for the numeral 'eight'

H0.1
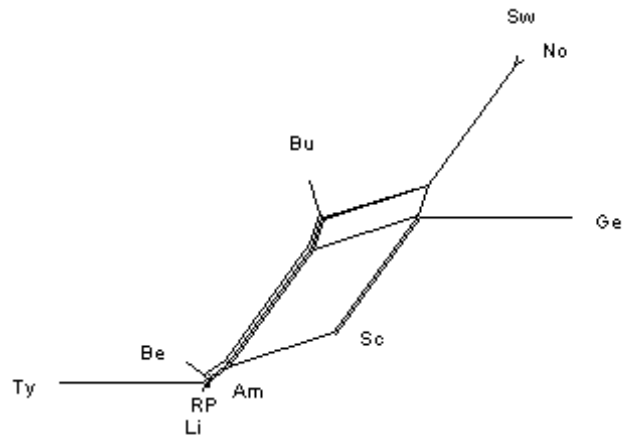
Figure C.5: NeighborNet diagram for the Connolly method for the numeral 'eight'
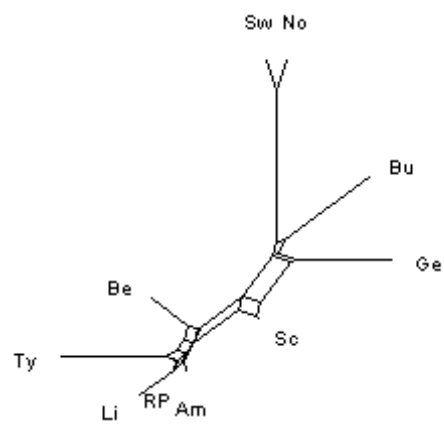
Figure C.6: NeighborNet diagram for the Almeida & Braun method for the numeral 'eight'

APPENDIX D

This appendix contains material relevant to chapter 4.

D.1   Experiment I

Statements:
1. We live in Ealing.
2. You remembered the lilies.
3. We arrived in a limo.
4. They are on the railings.
5. We were in yellow.
6. He is on the lilo.
7. You are feeling mellow.
8. We were lying.

Declarative questions:
1. He is on the lilo?
2. You remembered the lilies?
3. You live in Ealing?

Inversion questions:
1. May I lean on the railings?
2. May I leave the meal early?
3. Will you live in Ealing?

Wh questions:
1. Where is the manual?
2. When will you be in Ealing?
3. Why are we in a limo?

Coordination questions:
1. Are you growing limes or lemons?
2. Is his name Miller or Mailer?
3. Did you say mellow or yellow?
4. Do you live in Ealing or Reading?
5. Did he say lino or lilo?

Table D.1: IViE Sentence List data used (Grabe et al., 2001)

The Cinderella Passage (Grabe et al., 2001)

Once upon a time there was a girl called Cinderella. But everyone called her Cinders. Cinders lived with her mother and two stepsisters called Lily and Rosa. Lily and Rosa were very unfriendly and they were lazy girls. They spent all their time buying new clothes and going to parties. Poor Cinders had to wear all their old hand-me-downs! And she had to do the cleaning!

One day, a royal messenger came to announce a ball. The ball would be held at the Royal Palace, in honour of the Queen's only son, Prince William. Lily and Rosa thought this was divine. Prince William was gorgeous, and he was looking for a bride! They dreamed of wedding bells! When the evening of the ball arrived, Cinders had to help her sisters get ready. They were in a bad mood. They'd wanted to buy some new gowns, but their mother said that they had enough gowns. So they started shouting at Cinders. 'Find my jewels!' yelled one. 'Find my hat!' howled the other. They wanted hairbrushes, hairpins and hair spray.

When her sisters had gone, Cinders felt very down, and she cried. Suddenly, a voice said: 'Why are you crying, my dear?'. It was her fairy godmother! The girl poured her heart out: 'Lily and Rosa have it all!' she cried, 'even though they're awful, and fat, and they're dull! And I want to go to the ball, and meet Prince William!'

'You will, won't you?' laughed her fairy godmother. 'Go into the garden and find me a pumpkin'. Cinders went, and found a splendid pumpkin which the fairy changed into a dazzling carriage.

'Now bring me four white mice,' the godmother said. The girl went, and found one... two...three...four mice. The fairy godmother changed the mice into four lovely horses to pull the carriage.

Then the girl looked at her old rags. 'Oh dear!' she sighed. 'Where will I find something to wear? I don't have a gown!' 'Hmmm...' said the fairy : 'Let's see, what do you need? You'll need a ballgown... you need jewellery... you need shoes, and... something needs to be done about your hair. And would you like a blue gown or a green gown?'

For the third time, Cinders' godmother waved her magic wand. A ballgown, a robe and jewels appeared. And there were some elegant glass slippers. 'You

look wonderful,' her fairy godmother said, smiling. 'Just remember one thing - the magic only lasts until midnight!' And off Cinders went to the ball.

In the Royal Palace, everyone was amazed by the radiant girl in the beautiful ballgown. 'Who is she?' they asked. Prince William thought Cinders was the most beautiful girl he had ever seen. 'Have we met?' he asked. 'And may I have the honour of this dance?'

Prince William and Cinders danced for hours. Cinders was so glad that she failed to remember her fairy godmother's warning. Suddenly the clock chimed midnight! Cinders ran from the ballroom. 'Where are you going?' Prince William called. In her hurry, Cinders lost one of her slippers. The Prince wanted to find Cinderella, but he couldn't find the girl. 'I don't even know her name,' he sighed. But he held on to the slipper.

After the ball, the Prince was resolved to find the beauty who had stolen his heart. The glass slipper was his only clue. So he declared: 'The girl whose foot will fit this slipper shall be my wife'. And he began to search the kingdom.

Every girl in the land was willing to try on the slipper. But the slipper was always too small. When the Royal travellers arrived at Cinders' home, Lily and Rosa tried to squeeze their feet into the slipper. But it was no use; their feet were enormous! 'Do you have any other girls?' the Prince asked Cinders' mother. 'One more,' she replied. 'Oh no,' cried Lily and Rosa. 'She is much too busy!' But the Prince insisted that all girls must try the slipper.

Cinders was embarrassed. She didn't want the Prince to see her in her old apron. And her face was dirty! 'This is your daughter?' the Prince asked, amazed. But then Cinders tried on the glass slipper, and it fitted perfectly!

The Prince looked carefully at the girl's face, and he recognised her. 'It's you, my darling isn't it?' he yelled. 'Will you marry me?' Lily and Rosa were horrified. 'It was you at the ball, Cinders?' they asked. They couldn't believe it! Then Cinders married William, and they lived happily ever after.

I modified versions of the following scripts by Pauline Welby (original scripts available at http://www.icp.inpg.fr/˜welby/PAGES/software.html (Feb 2011) for use in the IViE analysis:

1. check.praat
2. draw-waveform-sgram-f0.praat
3. the set of three 'elbow' scripts
4. getvaluestones.praat
5. labeltones.praat
6. readin-files.praat
7. savesmallfiles.praat

Table D.2: Summary of scripts used

Category	Number of IPs analysed
Belfast statement rises (Sentences list)	74 IPs (6 male, 6 female)
Bel. stat. rises (Cinderella passage)	65 IPs (6 f., 2 m.)
Bel. question rises	92 IPs (6 m., 6 f.) (Sentences list only)
Cambridge stat. (rise)-falls	67 IPs (6 f., 5 m.) (Cinderella passage only)
Cam. q. rises	48 IPs (6 m., 6 f.) (Sentences list only)

Table D.3: Number of IPs analysed in Experiment I analysis of IViE data

Category	Question Breakdown
Bel. q. rises	33 declarative qs, 28 y/n inversion qs, 31 wh qs
Cam. q. rises	13 coordination qs, 23 declarative qs, 10 y/n inversion qs, 2 wh qs

Table D.4: Breakdown of the number of different question types analysed in the IViE analysis

Category	Mean alignment of L
Cam. stat. (rise)-falls	38 ms after onset of prenuclear vowel (s.d. 72)
Bel. stat. rises (Sentences list)	10 ms after offset of nuclear vowel (s.d. 48)
Bel. stat. rises (Cinderella passage)	11 ms after offset of nuclear vowel (s.d. 71)
Bel. q. rises	19 ms after offset of nuclear vowel (s.d. 40)
Cam. q. rises	6 ms before offset of nuclear vowel (s.d. 87)

Table D.5: Mean alignment values for L from the IViE analysis

Category	Mean alignment of H
Cam. stat. (rise)-falls	30 ms after onset of nuclear vowel (s.d. 70)
Bel. stat. rises (Sentences list)	110 ms after onset of postnuclear vowel (s.d. 60)
Bel. stat. rises (Cinderella passage)	101 ms after onset of postnuclear vowel (s.d. 53)
Bel. q rises	100 ms after onset of postnuclear vowel (s.d. 50)
Cambridge q rises	94 ms after onset of postnuclear vowel (s.d. 118)

Table D.6: Mean alignment values for H from the IViE analysis

Category	Mean $f0$ excursion (Hz)
Cam. stat.	18 (s.d. 15)
Belfast stat. (Sentences list)	19 (s.d. 14)
Bel. stat. (Cinderella passage)	15 (s.d. 10)
Bel. q.	23 (s.d. 16)
Cam. q.	30 (s.d. 15)

Table D.7: $f0$ excursion (Hz) between L and H: Mean values (mixture of male and female speakers)

Category	t test (Bonferroni corrected)
Cam. stat. vs. Cam. q.	t-3.2046, $p < 0.05$
Cam. stat. vs. Bel. q.	t -1.5351, n.s.
Cam. stat. vs. Bel. stat. (Sentences list)	t -0.3858, n.s.
Cam. stat. vs. Bel. stat. (Cinderella passage)	t 1.1373, n.s.
Cam. q. vs. Bel. q.	t 2.0832, n.s.
Cam. q. vs. Bel. stat. (Sentences)	t -3.2877, $p < 0.01$
Cam. q. vs. Bel. stat. (Cinderella)	t -4.9319, $p < 0.001$
Bel. q. vs. Bel. stat. (Sentences)	t -1.3949, n.s.
Bel. q. vs. Bel. stat. (Cinderella)	t -3.3399, $p < 0.05$
Bel. stat. (Sentences) vs. Bel. stat. (Cinderella)	t 1.9293, n.s.

Table D.8: $f0$ excursion: Bonferroni corrected post-hoc t tests (10 comparisons). Non-parametric tests returned very similar results.

Sentence type	Pearson product-moment correlation	Spearman's rank correlation rho
Belfast statements (passage)	t 1.1858, d.f. 59, n.s.	S 35117.89, n.s.
Belfast statements (sentences)	t 1.8517, d.f. 70, n.s.	S 33451.61, $p < 0.001$

Table D.9: Correlation between the alignment of L (measured from the offset of the nuclear vowel) and the alignment of H (measured from the onset of the postnuclear vowel)

D.2 Experiment II

Before conducting Experiment II, I filled in the forms complying with the Ethics procedures for Linguistics and English Language at the University of Edinburgh.

Initials	AG	CD	CR
Age	23	18	25
From	Lenzie, North Glasgow	South Glasgow	Northeast Glasgow
No. of years living there	23 years	18 years	17.5 years
Father from	North Glasgow	South Glasgow	Govan/Bearsden (in Glasgow)
Mother from	North Glasgow	South Glasgow	Buckie but lived in Glasgow for 35 years
Accent	typical	typical	fairly typical
Studying	Psychology	Law	PhD School of Community Health Sciences
Fluency in other languages	none	none	none
Tasks	Sentence reading and Map task	Sentence reading and Map task	Sentence reading only

Table D.10: Details of the three female Glasgow English speakers whose data I use in Experiment II

Initials	EL	PT	RB
Age	22	22	20
From	Southeast Belfast	East Belfast	South Belfast
No. of years living there	19	18	18
Father from	Southeast Belfast	Northeast Belfast	East Belfast
Mother from	Southeast Belfast	Northeast Belfast	County Tyrone
Accent	typical	typical	typical
Studying	Biology	Speech and Language Therapy	English literature and Spanish
Fluency in other languages	none	none	none
Tasks	Sentence reading and Map task	Sentence reading and Map task	Sentence reading and Map task

Table D.11: Details of the three female Belfast English speakers whose data I use in Experiment II

Initials	AR	CC	SK
Age	19	19	20
From	Polwarth	Black-hall/Drylaw	South Edinburgh
No. of years living there	19	19	20
Father from	Edinburgh	Edinburgh	Edinburgh
Mother from	Serbia	London	West Scotland
Accent[1]	typical	"apparently have a slight Inverness twang"	typical
Studying	English literature and history	German, Russian and Linguistics	medicine
Fluency in other languages	none	none	none
Tasks	Sentence reading and pilot Map task	Sentence reading and Map task	Sentence reading and pilot Map task

Table D.12: Details of the three female Edinburgh English speakers whose data I use in Experiment II

There was one key difference in my approach to segmentation from that in Turk et al. (2006). When delimiting vowels from surrounding voiceless plosives, I used Voice Onset Time (VOT) as my primary criterion, not the release of the plosive. The reason for this is that I designed my materials to make measurements of spectral tilt and peak amplitude on certain vowels. I wished to make these measurements on the voiced portion of vowels only. For consistency, I applied the VOT criterion to all vowels in my data where appropriate. To locate the onset and offset of voicing, I used the glottal pulses in the context of a continuous $f2$. For delimiting the beginning of a vowel after a voiced oral or nasal stop, I used the spike on the waveform indicating the stop release if present.
I did not attempt to segment post-vocalic /r/ from the preceding vowel if both were in the same syllable.

Table D.13: Extra notes about segmentation

1.	She walked on the path in Lala.
2.	The weather was bad in Vava.

Table D.14: Subset of 'Short tails' statements used in the analysis (5 repetitions)

1. She went through the land in Lala, through the field in Leely, and through the wood in Lolo.
2. The weather was bad in Vava, cold in Vovo, and sleety in Veevy.

Table D.15: Subset of 'Short tails' lists used in the analysis (5 repetitions)

1. Where is the path in Lala?
2. Why was the weather bad in Vava?

Table D.16: Subset of 'Short tails' wh questions used in the analysis (5 repetitions)

1. Did she walk on the path in Lala?
2. Was the weather bad in Vava?

Table D.17: Subset of 'Short tails' y/n questions used in the analysis (5 repetitions)

1. They stocked the vending machine.
2. They set the dining table.
3. They broke the rowing machine.
4. She liked the animator.
5. It's a War memorial.
6. It's an outdoor volleyball court.
7. She milked the dairy cattle.
8. She feared the gladiator.
9. The students loved the bungee jumping.
10. They spent time in the amusement arcade.

Table D.18: Subset of 'Long tails' statements used in the analysis

1. They fixed the vending machine, broke the rowing machine, and robbed the jewellery store.
2. They set the dining table, baked the gingerbread man, and played the monopoly game.
3. They broke the rowing machine, robbed the jewellery store, and ran to the railway station.
4. She liked the animator, feared the gladiator, and loved the gingerbread man.
5. She saw the War memorial, went on the roller coaster, and spent time in the amusement arcade.
6. The team played on the outdoor volleyball court, went bungee jumping, and sped around the roller coaster.
7. She milked the dairy cattle, picked the elderberries, and set the dining table.
8. She feared the gladiator, loved the bungee jumping, and sped around the roller coaster.
9. The students loved the bungee jumping, sped around the roller coaster, and spent time in the amusement arcade.
10. They spent time in the amusement arcade, sped around the roller coaster, and loved the bungee jumping.

Table D.19: Subset of 'Long tails' lists used in the analysis

1. When did they stock the vending machine?
2. When did they set the dining table?
3. How did they break the rowing machine?
4. Why did she like the animator?
5. When did she see the War memorial?
6. When did the team play on the outdoor volleyball court?
7. When did she milk the dairy cattle?
8. Why did she fear the gladiator?
9. Why did the students love the bungee jumping?
10. When did they spend time in the amusement arcade?

Table D.20: Subset of 'Long tails' wh questions used in the analysis

All of these sentences were preceded by 'Are you sure?'

1. Did they stock the vending machine?
2. Did they set the dining table?
3. Did they break the rowing machine?
4. Did she like the animator?
5. Did she see the War memorial?
6. Did the team play on the volleyball court?
7. Did she milk the dairy cattle?
8. Did she fear the gladiator?
9. Did the students love the bungee jumping?
10. Did they spend time in the amusement arcade?

Table D.21: Subset of 'Long tails' y/n questions used in the analysis

Speaker	'Tail'	statement	list	wh question	y/n question
AG	'Short'	6	5	10	10
AG	'Long'	10	5	10	10
CD	'Short'	4	5	10	8
CD	'Long'	2	3	7	5
CR	'Short'	2	2	7	8
CR	'Long'	6	4	10	7

Table D.22: Breakdown of usable utterances analysed in the Sentence reading task from the Glasgow English speakers

Speaker	'Tail'	statement	list	wh question	y/n question
EL	'Short'	9	9	10	9
EL	'Long'	10	10	10	10
PT	'Short'	10	9	10	6
PT	'Long'	9	9	9	10
RB	'Short'	10	9	9	3
RB	'Long'	10	9	10	9

Table D.23: Breakdown of usable utterances analysed in the Sentence reading task from the Belfast English speakers

Speaker	'Tail'	wh question
AR	'Short'	8
AR	'Long'	6
CC	'Short'	6
CC	'Long'	4
SK	'Short'	8
SK	'Long'	9

Table D.24: Breakdown of usable utterances analysed (wh questions only) in the Sentence reading task from the Edinburgh English speakers

'Short tails' Sentence type	Belfast	Glasgow
statement	71 ms	-66 ms
list	100 ms	88 ms
wh question	60 ms	-65 ms
y/n question	60 ms	13 ms

Table D.25: Mean absolute alignment values for L1 from the Begv1 ('Short tails'). A minus value indicates that the alignment of L1 occurred before Begv1.

'Long tails' Sentence type	Belfast	Glasgow
statement	73 ms	16 ms
list	119 ms	480 ms
wh question	38 ms	9 ms
y/n question	49 ms	129 ms

Table D.26: Mean absolute alignment values for L1 from Begv1 ('Long tails')

'Short tails' Sentence type	Belfast	Glasgow
statement	-64 ms	-214 ms
list	-48 ms	-54 ms
wh question	-66 ms	-208 ms
y/n question	-73 ms	-117 ms

Table D.27: Mean absolute alignment values for L1 from Endv1 ('Short tails') (for comparison with Experiment I)

'Long tails' Sentence type	Belfast	Glasgow
statement	-15 ms	-80 ms
list	23 ms	392 ms
wh question	-46 ms	-83 ms
y/n question	-41 ms	40 ms

Table D.28: Mean absolute alignment values for L1 from Endv1 ('Long tails') (for comparison with Experiment I)

'Short tails' Sentence type	Belfast	Glasgow
statement	0.573 (s.d. 0.42)	-0.471 (s.d. 0.651)
list	0.689 (s.d. 0.326)	0.603 (s.d. 0.838)
wh question	0.382 (s.d. 1.496)	-0.476 (s.d. 0.662)
y/n question	0.45 (s.d. 0.397)	0.156 (s.d. 1.063)

Table D.29: Mean proportional alignment values for L1 from Begv1 ('Short tails') (repeated for clarity from chapter 4)

'Long tails' Sentence type	Belfast	Glasgow
statement	0.886 (s.d. 0.587)	0.228 (s.d. 0.788)
list	1.3 (s.d. 1.093)	5.072 (s.d. 8.641)
wh question	0.414 (s.d. 0.899)	-0.033 (s.d. 1.327)
y/n question	0.525 (s.d. 0.88)	1.726 (s.d. 2.642)

Table D.30: Mean proportional alignment values for L1 from Begv1 ('Long tails') (repeated for clarity from chapter 4)

Main effect	Significance level
Statements vs. lists	$pmcmc < 0.001$
Wh questions vs. lists	$pmcmc < 0.001$
Wh questions vs. y/n questions	$pmcmc < 0.05$

Table D.31: Significant main effects, Alignment L1. In the Glasgow data alone, the results were extremely similar to the results for all the data pooled together. In all tables of statistical results, I report only statistically significant main effects and interactions to save space.

Main effect	Significance level
Statements vs. lists	pmcmc < 0.05
Statements vs. wh questions	pmcmc < 0.05
Wh questions vs. lists	pmcmc < 0.001

Table D.32: Significant main effects, Alignment L1, Belfast only

'Short tails' Sentence type	Belfast	Glasgow
statement	223 ms	-36 ms
list	219 ms	197 ms
wh question	195 ms	13 ms
y/n question	158 ms	85 ms

Table D.33: Mean absolute alignment values for H1 from Endv1 ('Short tails'). A minus value indicates that the alignment of H1 occurred before Endv1.

'Long tails' Sentence type	Belfast	Glasgow
statement	458 ms	125 ms
list	520 ms	359 ms
wh question	387 ms	170 ms
y/n question	362 ms	269 ms

Table D.34: Mean absolute alignment values for H1 from Endv1 ('Long tails')

'Short tails' Sentence type	Belfast	Glasgow
statement	0.82 (s.d. 0.138)	-0.182 (s.d. 0.596)
list	0.839 (s.d. 0.189)	0.732 (s.d. 0.494)
wh question	0.726 (s.d. 0.144)	0.039 (s.d. 0.478)
y/n question	0.629 (s.d. 0.121)	0.38 (s.d. 0.531)

Table D.35: Mean proportional alignment values for H1 from Endv1 as a proportion of the 'tail' ('Short tails') (repeated for clarity from chapter 4)

'Long tails' Sentence type	Belfast	Glasgow
statement	0.84 (s.d. 0.184)	0.235 (s.d. 0.215)
list	0.909 (s.d. 0.174)	0.651 (s.d. 0.426)
wh question	0.741 (s.d. (0.19)	0.332 (s.d. 0.304)
y/n question	0.723 (s.d. 0.211)	0.541 (s.d. 0.392)

Table D.36: Mean proportional alignment values for H from Endv1 as a proportion of the 'tail' ('Long tails') (repeated for clarity from chapter 4)

Sentence type	'Short'	'Long'
statement	0.729 (s.d. 0.186)	0.944 (s.d. 1.091)
list	0.788 (s.d. 0.248)	1.302 (s.d. 1.049)
wh question	0.628 (s.d. 0.162)	0.521 (s.d. 1.05)
y/n question	0.519 (s.d. 0.17)	0.23 (s.d. 0.971)

Table D.37: Mean proportional alignment values for H from beginning of final vowel as proportion of duration of final vowel-Belfast only. Values above 1 indicate that H1 was aligned beyond the IP-final vowel i.e. in the IP-final coda.

Main Effects and Interactions	Significance level
Variety	pmcmc < 0.01
Sentence type: wh questions vs. lists	pmcmc < 0.05
Variety*Tail	pmcmc < 0.01
Variety*Sentence type: statements vs. lists	pmcmc < 0.05
Variety*Sentence type: wh questions vs. y/n questions	pmcmc < 0.01
Variety*Tail*Sentence type: statements vs. lists	pmcmc < 0.05
Variety*Tail*Sentence type: statements vs. y/n questions	pmcmc < 0.05

Table D.38: Significant main effects and interactions, Alignment H1 as a proportion of the duration of the 'tail'

Main Effects	Significance level
Tail	pmcmc < 0.05
Sentence type: Statements vs. lists	pmcmc < 0.05
Wh questions vs. lists	pmcmc < 0.05

Table D.39: Significant main effects, Alignment H1 proportional, Glasgow

Main Effects and Interactions	Significance level
Sentence type: Statements vs. wh questions	pmcmc < 0.05
Statements vs. y/n questions	pmcmc < 0.01
Wh questions vs. lists	pmcmc = 0.001

Table D.40: Significant main effects, Alignment H1 proportional, Belfast

Main Effects and Interactions	Significance level
Sentence type: Statements vs. wh questions	pmcmc < 0.05
Statements vs. y/n questions	pmcmc < 0.001
Wh questions vs. lists	pmcmc < 0.001

Table D.41: Significant main effects, Alignment H1 proportional (beg. last vowel as proportion of duration of final vowel), Belfast

'Short tails' Sentence type	Belfast	Glasgow
statement	0.229 (s.d. 0.015)	0.22 (s.d. 0.012)
list	0.231 (s.d. 0.021)	0.228 (s.d. 0.021)
wh question	0.225 (s.d. 0.018)	0.216 (s.d. 0.022)
y/n question	0.22 (s.d. 0.016)	0.207 (s.d. 0.018)

Table D.42: Mean nuclear tone rate (duration/no. of sylls) ('Short tails') (repeated for clarity from chapter 4)

'Long tails' Sentence type	Belfast	Glasgow
statement	0.177 (s.d. 0.028)	0.182 (s.d. 0.031)
list	0.189 (s.d. 0.027)	0.193 (s.d. 0.031)
wh question	0.176 (s.d. 0.026)	0.177 (s.d. 0.022)
y/n question	0.166 (s.d. 0.016)	0.142 (s.d. 0.018)

Table D.43: Mean nuclear tone rate (duration/no. of sylls) ('Long tails') (repeated for clarity from chapter 4)

Main Effects	Significance level
Tail	pmcmc < 0.05
Sentence type: Statements vs. y/n questions	pmcmc < 0.01
Wh questions vs. y/n questions	pmcmc < 0.01

Table D.44: Significant main effects, 'Nuclear tone rate'

'Short tails' Sentence type	AG	CD	CR
statement	23	18	16
list	24	47	12
wh question	52	59	33
y/n question	49	63	20

Table D.45: Mean $f0$ excursion (Hz) ('Short tails', Glasgow)

'Short tails' Sentence type	EL	PT	RB
statement	48	36	43
list	43	27	54
wh question	75	55	74
y/n question	87	74	88

Table D.46: Mean $f0$ excursion (Hz) ('Short tails', Belfast)

'Long tails' Sentence type	AG	CD	CR
statement	38	46	16
list	41	47	9
wh question	54	59	24
y/n question	66	62	29

Table D.47: Mean $f0$ excursion (Hz) ('Long tails', Glasgow)

'Long tails' Sentence type	EL	PT	RB
statement	62	36	42
list	49	40	43
wh question	105	81	80
y/n question	94	90	86

Table D.48: Mean $f0$ excursion (Hz) ('Long tails', Belfast)

Speaker	Main Effects and interactions	p value
AG	Sentence type: Statements vs. wh questions	pmcmc < 0.05
AG	Statements vs. y/n questions	pmcmc < 0.001
CR	Statements vs. y/n questions	pmcmc < 0.05
CR	Wh questions vs. lists	pmcmc < 0.05
CR	Tail*Sentence type: wh questions vs. y/n questions	pmcmc < 0.05
EL	Tail	pmcmc < 0.05
EL	Statements vs. lists	pmcmc < 0.01
EL	Statements vs. wh questions	pmcmc < 0.001
EL	Statements vs. y/n questions	pmcmc < 0.001
EL	Wh questions vs. lists	pmcmc < 0.001
EL	Tail*Sentence type: Wh questions vs. lists	pmcmc < 0.05
EL	Tail*Sentence type: wh questions vs. y/n questions	pmcmc < 0.05
PT	Statements vs. wh questions	pmcmc < 0.001
PT	Statements vs. y/n questions	pmcmc < 0.001
PT	Wh questions vs. lists	pmcmc < 0.001
PT	Tail*Sentence type: Statements vs. wh questions	pmcmc < 0.05
RB	Statements vs. wh questions	pmcmc < 0.01
RB	Statements vs. y/n questions	pmcmc < 0.001
RB	Wh questions vs. lists	pmcmc < 0.01
RB	Tail*Sentence type: Wh questions vs. lists	pmcmc < 0.01

Table D.49: Significant main effects and interactions, $f0$ excursion. Speaker CD was the only speaker for whom there were no differences at all.

1. Get the mean $f0$ of each utterance of a given speaker in logHz.
2. Get the duration of each utterance[2].
3. Multiply the mean $f0$ of each utterance in logHz by the utterance's duration (individual utterance weighted means).
4. Get the sum of the individual utterance durations (A).
5. Get the sum of all of individual utterance weighted means (B).
6. Divide B by A.

Table D.50: Calculation of speaker's weighted average for scaling normalisation

---

[2]The reason for including duration in the calculation of the weighted average of $f0$ is to take account of the fact that different utterances had different durations. Including duration allows longer utterances to contribute more to the speaker's weighted average than shorter utterances.

> 1. Divide the logHz value by $\log_{10}(2)$ to get a value in octaves.
> 2. Multiply this by 12 to convert to semitones.

Table D.51: Procedure for converting logHz to semitones (re 1 Hz) (http://www.mail-archive.com/praat-users@yahoogroups.co.uk/msg00356.html (Dec 2010), comment is from Daniel Hirst 29 March 2010)

'Short tails' Sentence type	Belfast	Glasgow
statement	-2.778 (s.d. 1.051)	-1.649 (s.d. 1.272)
list	-2.568 (s.d. 1.381)	-2.568 (s.d. 1.368)
wh question	-2.719 (s.d. 1.202)	-1.917 (s.d. 2.426)
y/n question	-2.352 (s.d. 1.458)	-2.434 (s.d. 1.211)

Table D.52: Mean normalised scaling of L1 in semitones ('Short tails') (repeated for clarity from chapter 4)

Long tails Sentence type	Belfast	Glasgow
statement	-2.274 (s.d. 1.188)	-1.202 (s.d. 0.96)
list	-1.231 (s.d. 0.757)	-1.694 (s.d. 1)
wh question	-2.394 (s.d. 1.372)	-1.364 (s.d. 1.385)
y/n question	-1.692 (s.d. 1.189)	-1.688 (s.d. 1.415)

Table D.53: Mean normalised scaling of L1 in semitones ('Long tails') (repeated for clarity from chapter 4)

> There was a significant main effect of Tail $p < 0.001$. Considering the Glasgow and Belfast data separately, the results were extremely similar.

Table D.54: Significant main effect, L1 scaling

'Short tails' Sentence type	Belfast	Glasgow
statement	0.964 (s.d. 0.723)	0.136 (s.d. 1.161)
list	1.015 (s.d. 2.075)	0.016 (s.d. 0.899)
wh question	2.471 (s.d. 0.998)	1.991 (s.d. 1.412)
y/n question	4.141 (s.d. 1.391)	1.124 (s.d. 1.697)

Table D.55: Mean normalised scaling of H1 in semitones ('Short tails') (repeated for clarity from chapter 4)

'Long tails' Sentence type	Belfast	Glasgow
statement	1.808 (s.d. 2.718)	1.242 (s.d. 1.46)
list	2.369 (s.d. 1.14)	0.355 (s.d. 1.437)
wh question	4.551 (s.d. 1.075)	2.188 (s.d. 1.72)
y/n question	5.088 (s.d. 1.346)	2.514 (s.d. 1.879)

Table D.56: Mean normalised scaling of H1 in semitones 'Long tails' (repeated for clarity from chapter 4)

Main Effects	Significance level
Sentence type: Statements vs. wh questions	pmcmc < 0.05
Statements vs. y/n questions	pmcmc < 0.05
Wh questions vs. lists	pmcmc < 0.01

Table D.57: Significant main effects, H1 Scaling, Glasgow

Main Effects and Interactions	Significance level
Sentence type: Statements vs. wh questions	pmcmc < 0.001
Statements vs. y/n questions	pmcmc < 0.001
Wh questions vs. lists	pmcmc < 0.001
Tail*Sentence type: Statements vs. wh questions	pmcmc < 0.05

Table D.58: Significant main effects and interactions, H1 Scaling, Belfast

'Short tails' Sentence type	Belfast	Glasgow
statement	-1.534 (s.d. 1.332)	-1.96 (s.d. 1.367)
list	-1.552 (s.d. 1.454)	-2.383 (s.d. 1.273)
wh question	-0.876 (s.d. 1.956)	-0.948 (s.d. 1.856)
y/n question	0.903 (s.d. 2.706)	-1.738 (s.d. 1.745)

Table D.59: Mean normalised scaling of onset of final vowel in semitones ('Short tails') (repeated for clarity from chapter 4)

'Long tails' Sentence type	Belfast	Glasgow
statement	1.574 (s.d. 1.672)	-0.159 (s.d. 1.112)
list	0.623 (s.d. 1.194)	-1.121 (s.d. 0.767)
wh question	3.182 (s.d. 2.521)	-1.877 (s.d. 2.787)
y/n question	1.992 (s.d. 4.313)	-0.546 (s.d. 1.361)

Table D.60: Mean normalised scaling of onset of final vowel in semitones ('Long tails') (repeated for clarity from chapter 4)

Main Effects and Interactions	Significance level
Variety	pmcmc < 0.05
Sentence type: Statements vs. wh questions	pmcmc < 0.01
Statements vs. y/n questions	pmcmc < 0.01

Table D.61: Significant main effects and interactions, Scaling of onset of final vowel

'Tail'	Edinburgh	Glasgow
'Short'	-0.525 (0.263)	-0.476 (0.662)
'Long'	-0.42 (0.4)	-0.033 (1.327)

Table D.62: Mean proportional alignment of L1 in the Edinburgh and Glasgow wh questions

'Tail'	Edinburgh	Glasgow
'Short'	-0.293 (s.d. 0.193)	0.039 (s.d. 0.478)
'Long'	0.035 (s.d. 0.119)	0.332 (s.d. 0.304)

Table D.63: Mean proportional alignment of H1 (measured from Endv1 as a proportion of the 'tail') in the Edinburgh and Glasgow wh questions

Main Effects	Significance level
Variety	pmcmc < 0.05
Tail	pmcmc = 0.001

Table D.64: Significant main effects, alignment H1, Edinburgh and Glasgow wh questions

'Tail'	Edinburgh	Glasgow
'Short'	-1.035 (s.d. 2.325)	-1.917 (s.d. 1.547)
'Long'	-0.665 (s.d. 1.385)	-1.364 (s.d. 2.426)

Table D.65: Mean normalised scaling of L1 in the Edinburgh and Glasgow wh questions in semitones

'Tail'	Edinburgh	Glasgow
'Short'	1.029 (s.d. 0.494)	1.991 (s.d. 1.412)
'Long'	2.111 (s.d 0.735)	2.188 (s.d. 1.72)

Table D.66: Mean normalised scaling of H1 in the Edinburgh and Glasgow wh questions in semitones

Speaker	Role	'Tail'	non-question	question
AG	Instruction Giver	'Short'	8	4
AG	Instruction Giver	'Long'	8	5
AG	Instruction Follower	'Short'	1	1
AG	Instruction Follower	'Long'	3	1
CD	Instruction Giver	'Short'	3	4
CD	Instruction Giver	'Long'	8	4

Table D.67: Breakdown of usable utterances analysed in the Glasgow Map tasks

Speaker	Role	'Tail'	non-question	question
EL	Instruction Giver	'Short'	3	1
EL	Instruction Giver	'Long'	8	3
PT	Instruction Follower	'Short'	13	2
PT	Instruction Follower	'Long'	9	5
RB	Instruction Giver	'Short'	19	8
RB	Instruction Giver	'Long'	19	4

Table D.68: Breakdown of usable utterances analysed in the Belfast Map tasks

Amusement arcade

Gladiator

Valley

Merry go round

Roller coaster

Amusement arcade

Lemmings

Windmill

Lilies

Bungee jumping

X Finish

Bowling alley

Arrow
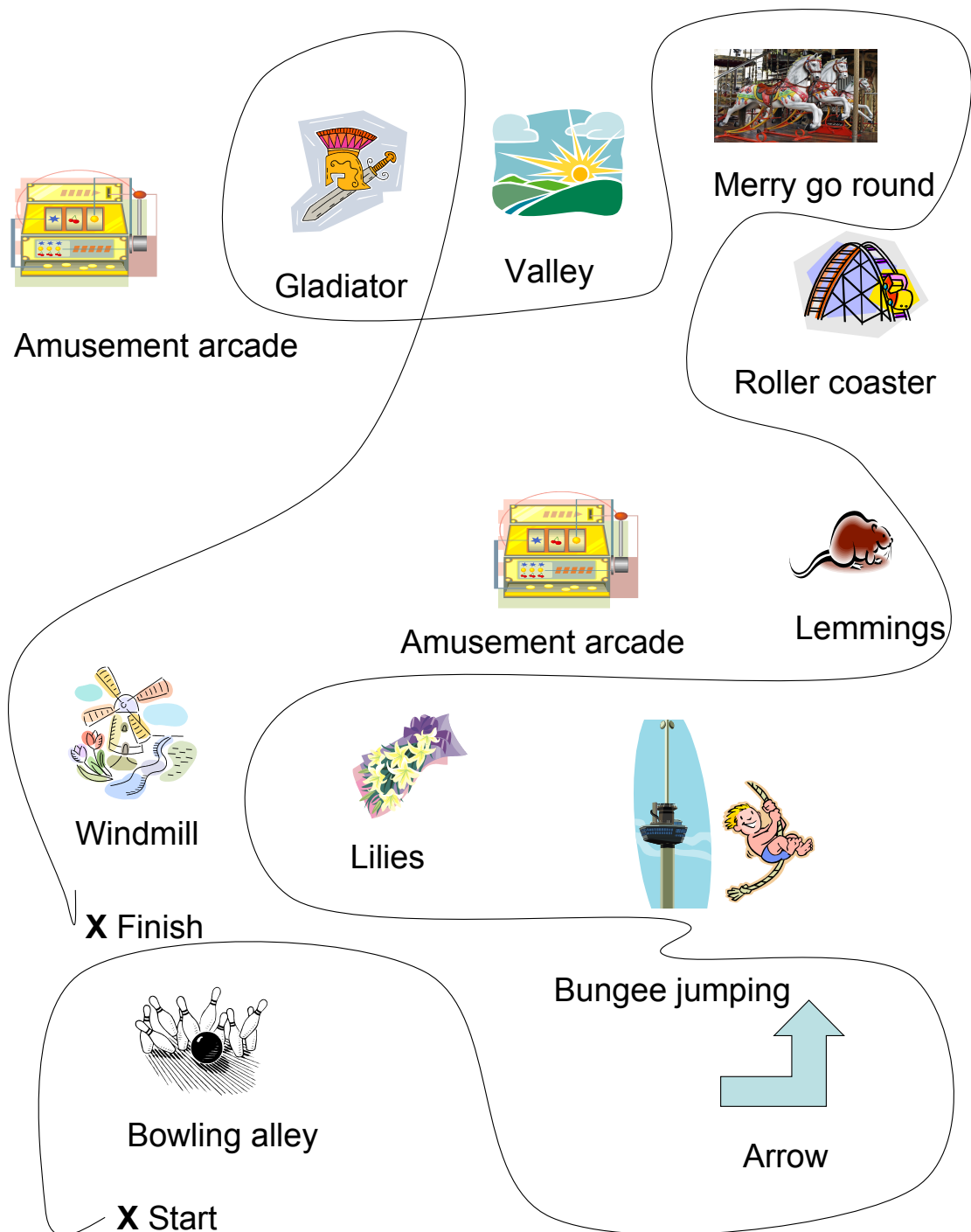
X Start

Map 1: Instruction Giver

Figure D.1: Map 1: Instruction giver

Amusement arcade

Roller coaster

Valley

Roller coaster

Gold mine

Vending machine

Windmill

Lilies

Bungee jumping

Wildlife reserve

Arrow
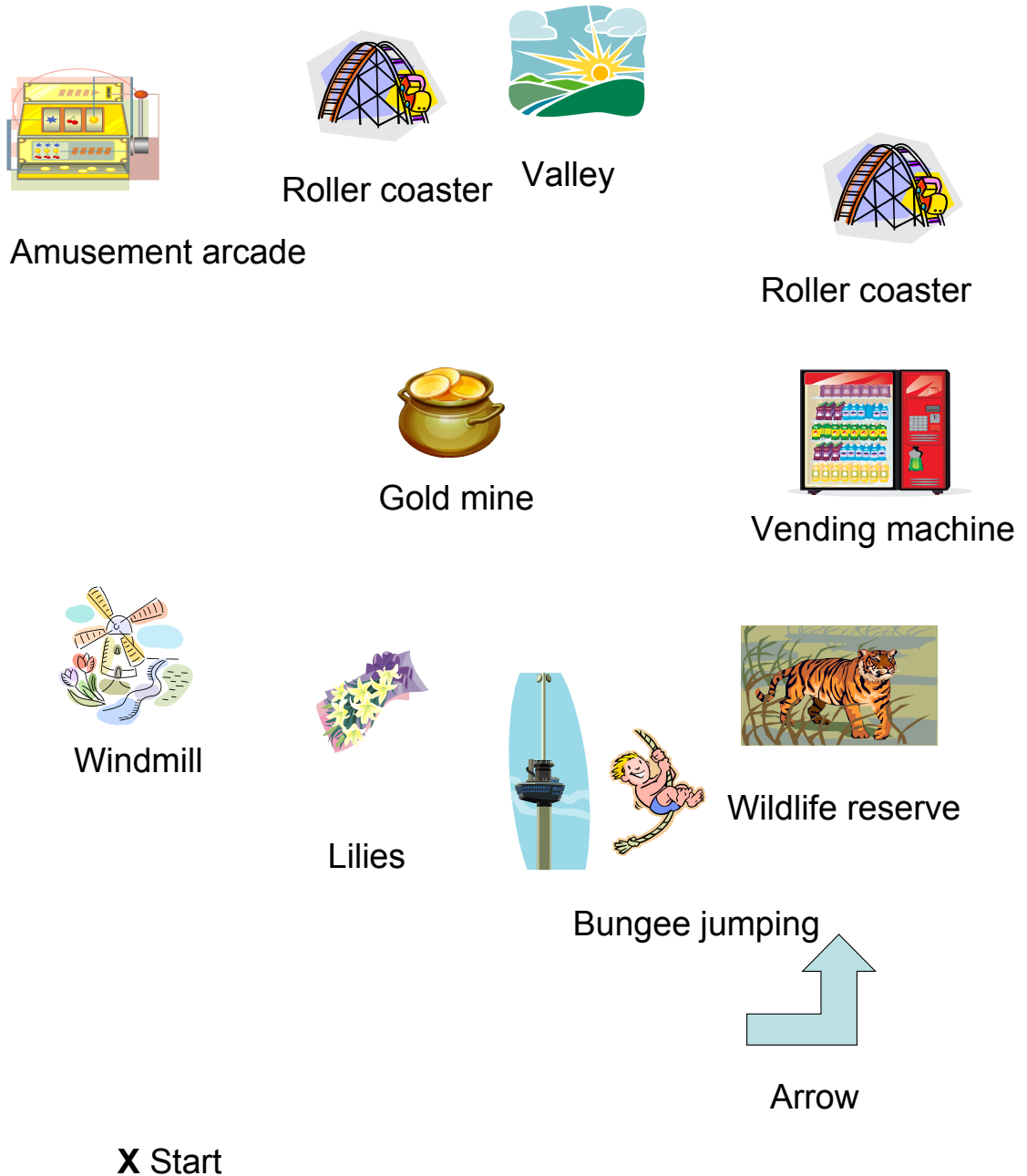
X Start

Map 1: Follower

Figure D.2: Map 1: Instruction follower
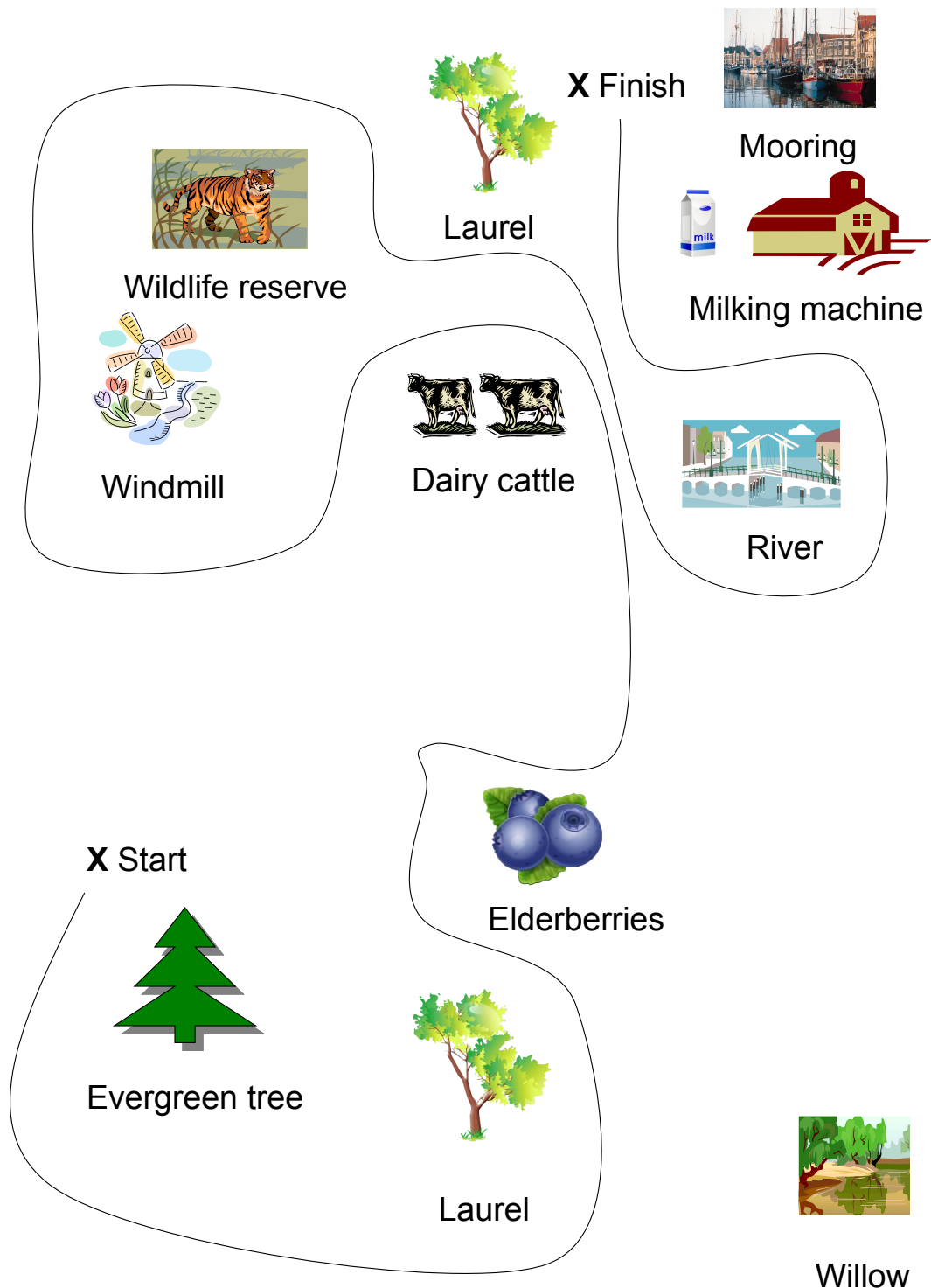
Map 2: Instruction Giver

Figure D.3: Map 2: Instruction giver

Laurel

Mooring

Merry go round

Windmill

Dairy cattle

River

X Start
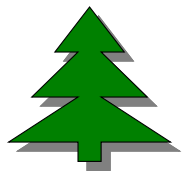
Elderberries

Loganberries

Evergreen tree

Willow

Map 2: Follower

Figure D.4: Map 2: Instruction follower

X Start

Animator

Laneway

Railway station

Lorry

Volleyball court

Lilies

Wedding chapel

Volleyball court

War memorial

X Finish

Runway

The Underbelly
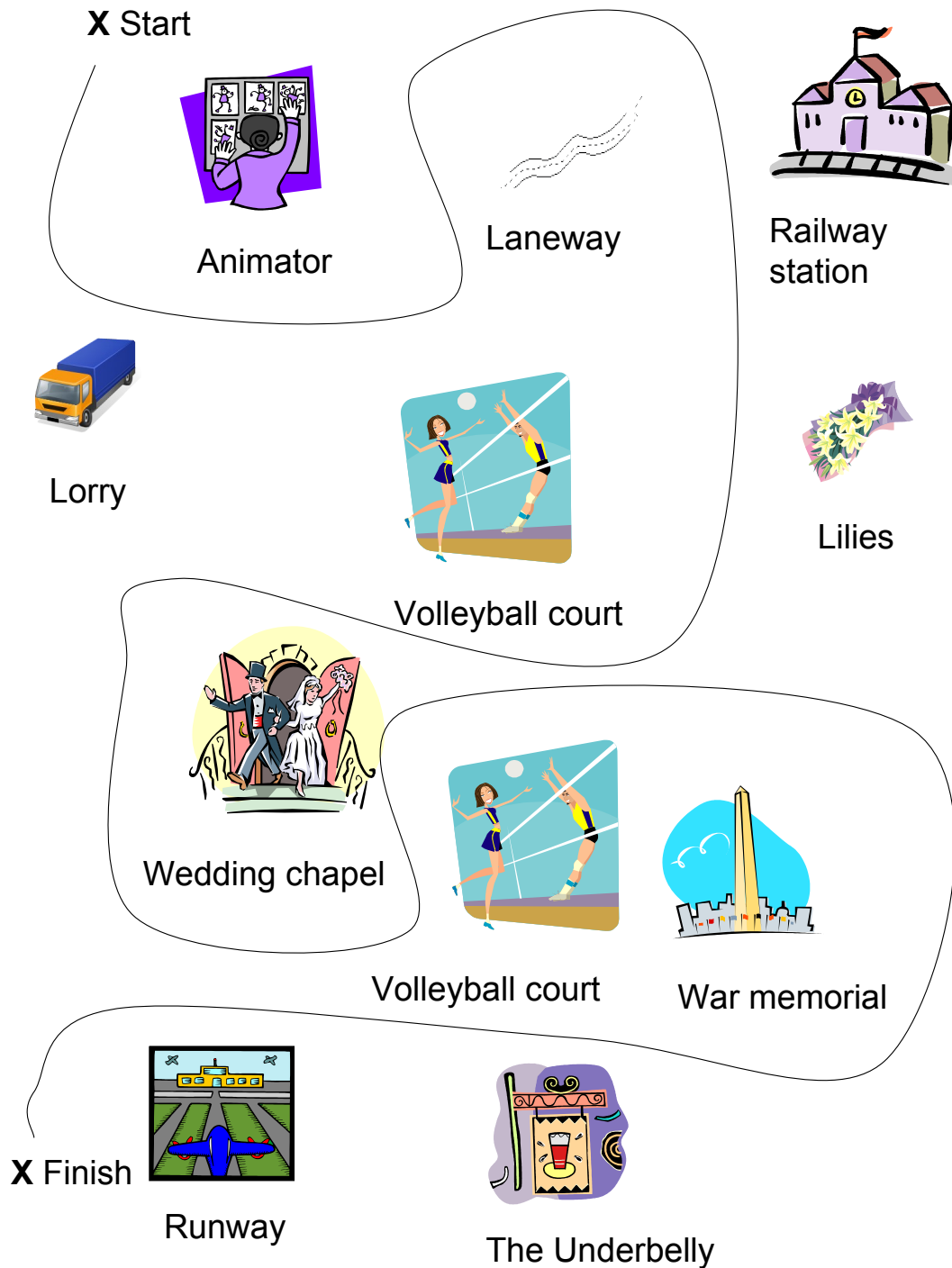
Map 4: Instruction Giver

Figure D.5: Map 4: Instruction giver. Note that Map 3 was not included in the subset of data referred to in chapter 4

X Start

Animator

Railway station

Lorry

Elderberries

Roller coaster

Roses

Merry go round

Volleyball court
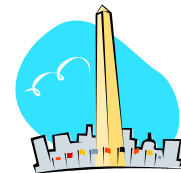
War memorial
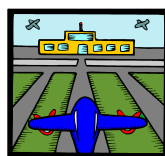
Runway

Map 4: Follower

Figure D.6: Map 4: Instruction follower

'Tail'	Contour type	Utterance type	Alignment L1
'Short'	No final low	Non q	0.127 (s.d. 1.474)
'Long'	No final low	Non q	0.633 (s.d. 1.046)
'Short'	Final low	Non q	NA
'Long'	Final low	Non q	-0.779 (s.d. 0.297)
'Short'	No final low	Q	0.553 (s.d. 1.514)
'Long'	No final low	Q	3.186 (s.d. 3.642)
'Short'	Final low	Q	-1.096 (s.d. 1.008)
'Long'	Final low	Q	-0.627 (s.d. 0.887)

Table D.69: Mean proportional alignment of L1 (Glasgow Map task). NA means that there were no 'Short' non question utterances with a final low in the data.

'Tail'	Contour type	Utterance type	Alignment H1
'Short'	No final low	Non q	0.801 (s.d. 0.264)
'Long'	No final low	Non q	0.888 (s.d. 0.132)
'Short'	Final low	Non q	NA
'Long'	Final low	Non q	0.533 (s.d. 0.367)
'Short'	No final low	Q	0.852 (s.d. 0.297)
'Long'	No final low	Q	0.853 (s.d. 0.236)
'Short'	Final low	Q	0.458 (s.d. 0.22)
'Long'	Final low	Q	0.247 (s.d. 0.096)

Table D.70: Mean proportional alignment of H1 (Glasgow Map task)

Main Effect	Significance level
Contour type	$pmcmc < 0.05$

Table D.71: Significant main effect, alignment H1 (Glasgow Map task)

'Tail'	Utterance type	Nuclear tone rate
'Short'	Non q	0.199 (s.d. 0.031)
'Long'	Non q	0.172 (s.d. 0.031)
'Short'	Q	0.21 (s.d. 0.05)
'Long'	Q	0.18 (s.d. 0.019)

Table D.72: Mean 'Nuclear tone rate' (Glasgow Map task)

Main Effect	Significance level
Tail	pmcmc < 0.001

Table D.73: Significant main effect, 'Nuclear tone rate' (Glasgow Map task)

'Tail'	Contour type	Utterance type	AG	CD
'Short'	No final low	Non q	44 (s.d. 25)	69 (s.d. 38)
'Long'	No final low	Non q	94 (s.d. 59)	94 (s.d. 37)
'Short	Final Low	Non q	NA	NA
'Long'	Final low	Non q	59 (s.d. 10)	NA
'Short'	No final low	Q	62 (s.d. 0.707)	77 (s.d. 45)
'Long'	No final low	Q	63 (s.d. NA)	73 (s.d. 23)
'Short'	Final low	Q	52 (s.d. 24)	NA
'Long'	Final low	Q	NA	NA

Table D.74: Mean $f0$ excursion (Glasgow Map task)

'Tail'	Contour type	Utterance type	Scaling L1
'Short'	No final low	Non q	-1.035 (s.d. 0.672)
'Long'	No final low	Non q	-1.29 (s.d. 1.63)
'Short'	Final low	Non q	NA
'Long'	Final low	Non q	-0.514 (s.d. 0.304)
'Short'	No final low	Q	-1.29 (s.d. 1.345)
'Long'	No final low	Q	-1.02 (s.d. 1.79)
'Short'	Final low	Q	-0.405 (s.d. 0.45)
'Long'	Final low	Q	-0.604 (s.d. NA)

Table D.75: Mean scaling of L1 in semitones (Glasgow Map task)

'Tail'	Contour type	Utterance type	Scaling H1
'Short'	No final low	Non q	3.355 (s.d. 2.571)
'Long'	No final low	Non q	5.122 (s.d. 2.754)
'Short'	Final low	Non q	NA
'Long'	Final low	Non q	1.748 (s.d. 1.594)
'Short'	No final low	Q	3.886 (s.d. 2.023)
'Long'	No final low	Q	3.309 (s.d. 1.811)
'Short'	Final low	Q	3.356 (s.d. 1.207)
'Long'	Final low	Q	3.861 (s.d. 0.937)

Table D.76: Mean scaling of H1 in semitones (Glasgow Map task)

'Tail'	Utterance type	Alignment L1
'Short'	Non q	0.203 (s.d. 1.63)
'Long'	Non q	1.005 (s.d. 1.7)
'Short'	Q	0.982 (s.d. 0.279)
'Long'	Q	0.551 (s.d. 0.716)

Table D.77: Mean proportional alignment of L1 (Belfast Map task)

'Tail'	Utterance type	Alignment H1
'Short'	Non q	0.753 (s.d. 0.234)
'Long'	Non q	0.831 (s.d. 0.189)
'Short'	Q	0.815 (s.d. 0.239)
'Long'	Q	0.816 (s.d. 0.23)

Table D.78: Mean proportional alignment of H1 (Belfast Map task)

'Tail'	Utterance type	Nuclear tone rate
'Short'	Non q	0.229 (s.d. 0.035)
'Long'	Non q	0.171 (s.d. 0.028)
'Short'	Q	0.207 (s.d. 0.055)
'Long'	Q	0.174 (s.d. 0.022)

Table D.79: Mean 'Nuclear tone rate' (Belfast Map task)

'Tail'	Utterance type	EL	PT	RB
'Short'	Non q	55 (s.d. 13)	56 (s.d. 30)	60 (s.d. 24)
'Long'	Non q	52 (s.d. 9)	48 (s.d. 14)	68 (s.d. 23)
'Short'	Q	37 (s.d. NA)	NA	62 (s.d. 17)
'Long'	Q	81 (s.d. 7)	101 (s.d. 9)	58 (s.d. 40)

Table D.80: Mean $f0$ excursion (Belfast Map task)

Speaker	Main Effect	p value
PT	Utterance type	$p = 0.001$

Table D.81: Significant main effect, $f0$ excursion (Belfast Map task)

'Tail'	Non-question	Question
'Short tails'	-1.656 (s.d. 1.279)	-1.574 (s.d. 0.677)
'Long tails'	-1.459 (s.d. 1.126)	-0.738 (s.d. 1.042)

Table D.82: Mean normalised scaling of L1 in semitones (Belfast Map task)

Main Effect	p value
Utterance type	$pmcmc < 0.05$

Table D.83: Significant main effect, L1 scaling (Belfast Map task)

'Tail'	Non-question	Question
'Short tails'	3.072 (s.d. 2.205)	3.503 (s.d. 1.484)
'Long tails'	3.44 (s.d. 1.59)	5.526 (s.d. 1.219)

Table D.84: Mean normalised scaling of H1 semitones (Belfast Map task)

Main Effect	p value
Utterance type	$pmcmc < 0.01$

Table D.85: Significant main effect, H1 scaling (Belfast Map task)

APPENDIX E

This appendix contains material relevant to chapter 5.

Abbreviations to figures E.1 and E.2: bel (Belfast), gl (Glasgow), st (statement). The top matrix is of L1 alignment, the second from top is of H1 alignment, the third is of L1 scaling, the fourth is of H1 scaling and the bottom matrix is of the overall intonational distance score.

L1 alignment

	bel_short_st	bel_short_list	bel_short_wh	bel_short_yn	gl_short_st	gl_short_list	gl_short_wh	gl_short_yn	bel_long_st	bel_long_list	bel_long_wh
bel_short_st	0										
bel_short_list	0	0									
bel_short_wh	0	0	0								
bel_short_yn	0	0	0	0							
gl_short_st	2	3	2	2	0						
gl_short_list	0	0	0	0	2	0					
gl_short_wh	2	3	2	2	0	2	0				
gl_short_yn	1	1	0	0	1	1	1	0			
bel_long_st	0	0	0	0	2	0	2	1	0		
bel_long_list	0	0	1	1	3	0	3	2	0	0	
bel_long_wh	0	1	0	0	2	0	2	0	0	1	0
bel_long_yn	0	1	0	0	2	0	2	0	0	1	0
gl_long_st	1	1	0	0	1	1	1	0	1	2	0
gl_long_list	8	7	8	8	9	7	9	9	8	6	8
gl_long_wh	1	1	1	1	1	1	1	0	1	2	0
gl_long_yn	1	0	1	1	3	0	3	2	1	0	1

H1 alignment

	bel_short_st	bel_short_list	bel_short_wh	bel_short_yn	gl_short_st	gl_short_list	gl_short_wh	gl_short_yn	bel_long_st	bel_long_list	bel_long_wh
bel_short_st	0										
bel_short_list	0	0									
bel_short_wh	0	0	0								
bel_short_yn	1	1	0	0							
gl_short_st	5	5	4	3	0						
gl_short_list	0	0	0	0	4	0					
gl_short_wh	4	4	3	2	0	3	0				
gl_short_yn	2	2	2	1	2	2	1	0			
bel_long_st	4	4	5	6	9	5	8	7	0		
bel_long_list	5	6	6	7	11	6	10	8	1	0	
bel_long_wh	3	3	2	4	8	3	7	6	1	2	0
bel_long_yn	2	2	2	4	7	3	6	5	1	3	0
gl_long_st	1	1	1	0	3	1	2	0	6	7	5
gl_long_list	2	2	3	4	7	3	6	5	1	3	0
gl_long_wh	1	0	0	0	4	0	2	1	5	7	4
gl_long_yn	0	1	1	2	6	1	5	3	3	5	2

L1 scaling

	bel_short_st	bel_short_list	bel_short_wh	bel_short_yn	gl_short_st	gl_short_list	gl_short_wh	gl_short_yn	bel_long_st	bel_long_list	bel_long_wh
bel_short_st	0										
bel_short_list	0	0									
bel_short_wh	0	0	0								
bel_short_yn	0	0	0	0							
gl_short_st	1	0	1	0	0						
gl_short_list	0	0	0	0	0	0					
gl_short_wh	0	0	0	0	0	0	0				
gl_short_yn	0	0	0	0	0	0	0	0			
bel_long_st	0	0	0	0	0	0	0	0	0		
bel_long_list	1	1	1	1	0	1	0	1	1	0	
bel_long_wh	0	0	0	0	0	0	0	0	0	1	
bel_long_yn	1	0	1	0	0	0	0	0	0	0	0
gl_long_st	1	1	1	1	0	1	0	1	1	0	0
gl_long_list	1	0	1	0	0	0	0	0	0	0	1
gl_long_wh	1	1	1	0	0	1	0	1	0	0	0
gl_long_yn	1	0	1	0	0	0	0	0	0	0	1

H1 scaling

	bel_short_st	bel_short_list	bel_short_wh	bel_short_yn	gl_short_st	gl_short_list	gl_short_wh	gl_short_yn	bel_long_st	bel_long_list	bel_long_wh
bel_short_st	0										
bel_short_list	0	0									
bel_short_wh	1	1	0								
bel_short_yn	4	4	1	0							
gl_short_st	0	0	2	4	0						
gl_short_list	0	0	2	4	0	0					
gl_short_wh	1	0	0	2	1	1	0				
gl_short_yn	0	0	1	3	0	1	0	0			
bel_long_st	0	0	0	2	1	1	0	0	0		
bel_long_list	2	2	0	1	2	2	0	1	0	0	
bel_long_wh	4	4	2	0	4	0	2	3	2	2	0
bel_long_yn	5	5	2	0	4	5	3	3	3	2	0
gl_long_st	0	0	1	2	1	1	0	0	0	1	3
gl_long_list	0	0	2	3	0	0	1	0	1	1	4
gl_long_wh	2	2	0	1	2	2	0	1	0	0	2
gl_long_yn	2	2	0	1	2	2	0	1	0	0	2

Overall Distance

	bel_short_st	bel_short_list	bel_short_wh	bel_short_yn	gl_short_st	gl_short_list	gl_short_wh	gl_short_yn	bel_long_st	bel_long_list	bel_long_wh
bel_short_st	0										
bel_short_list	0	0									
bel_short_wh	1	1	0								
bel_short_yn	5	5	1	0							
gl_short_st	8	8	9	9	0						
gl_short_list	0	0	2	4	6	0					
gl_short_wh	7	7	5	6	1	6	0				
gl_short_yn	3	3	3	4	3	4	2	0			
bel_long_st	4	4	5	8	12	6	10	8	0		
bel_long_list	8	9	8	10	16	9	13	12	2	0	
bel_long_wh	7	8	4	4	14	3	11	9	3	6	0
bel_long_yn	8	8	5	4	13	8	11	8	4	6	0
gl_long_st	3	3	3	3	5	4	3	1	8	10	8
gl_long_list	11	9	14	15	16	10	16	14	10	10	13
gl_long_wh	5	4	2	2	7	4	3	3	6	9	6
gl_long_yn	4	3	3	4	11	3	8	6	4	5	6

Figure E.1: Distance matrices for overall intonational distance measure (part 1)

bel_long_yn gl_long_st gl_long_list gl_long_wh gl_long_yn

```
0
0    0
8    9    0
0    0    9    0
1    2    6    2    0
```

bel_long_yn gl_long_st gl_long_list gl_long_wh gl_long_yn

```
0
4    0
0    4    0
3    0    3    0
1    2    1    1    0
```

bel_long_yn gl_long_st gl_long_list gl_long_wh gl_long_yn

```
0
0    0
0    0    0
0    0    0    0
0    0    0    0    0
```

bel_long_yn gl_long_st gl_long_list gl_long_wh gl_long_yn

```
0
3    0
4    0    0
2    0    1    0
2    1    2    0    0
```

bel_long_yn gl_long_st gl_long_list gl_long_wh gl_long_yn

```
0
7    0
12   13   0
5    0    13   0
4    5    9    3    0
```

Figure E.2: Distance matrices for overall intonational distance measure (part 2)

References

Adams, G., Barry, M. and Tilling, P. (1985), "The Tape-recorded Survey of Hiberno-English speech: A reappraisal of the techniques of traditional dialect geography", *in* J. Kirk, S. Sanderson and J. Widdowson, eds, *Studies in Linguistic Geography, the dialects of Britain and Ireland*, Croom Helm, London, pp. 67–80.

Akinlabi, A. and Liberman, M. (2001), "Tonal complexes and tonal alignment", *in Proceedings of the North East Linguistics Society (NELS)*, Vol. 31, pp. 1–20. **URL:** *http://www.ling.upenn.edu/˜myl/ (Feb 2011)*

Almeida, A. and Braun, A. (1986), "'Richtig' und 'falsch' in phonetischer Transkription: Vorschläge zum Vergleich von Transkriptionen mit Beispielen aus deutschen Dialekten", *Zeitschrift für Dialektologie und Linguistik*, LIII(2), pp.158–72.

Ambrazaitis, G. (2008), "On final rises and fall-rises in German and Swedish", *in Proceedings, FONETIK 2008*, Department of Linguistics, University of Gothenburg, pp. 81–4.

Anderson, A. H., Brown, M., Bard, E. G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S. and Weinert, R. (1991), "The HCRC Map Task Corpus", *Language and Speech*, 34(4), pp.351–66.

Arvaniti, A. and Garding, G. (2007), "Dialectal variation in the rising accents of American English", *in* J. Hualde and J. Cole, eds, *Papers in Laboratory Phonology IX*, Mouton de Gruyter, Berlin, pp. 547–76.

Arvaniti, A. and Ladd, D. R. (2009), "Greek wh-questions and the phonology of intonation", *Phonology*, 26, pp.43–74.

Arvaniti, A., Ladd, D. R. and Mennen, I. (1998), "Stability of tonal alignment: The case of Greek prenuclear accents", *Journal of Phonetics*, 26, pp.3–25.

Arvaniti, A., Ladd, D. R. and Mennen, I. (2000), "What is a starred tone? Evidence from Greek", *in* M. Broe and J. B. Pierrehumbert, eds, *Papers in Laboratory Phonology V*, CUP, Cambridge, pp. 119–31.

Arvaniti, A., Ladd, D. R. and Mennen, I. (2006), "Phonetic effects of focus and 'tonal crowding' in intonation: Evidence from Greek polar questions", *Speech Communication*, 48, pp.667–96.

Asu, E. L. (2006), "Tonal alignment in two varieties of Estonian", *in* G. Bruce and M. Horne, eds, *Nordic Prosody: Proceedings of the IXth Conference, Lund*, Peter Lang, Frankfurt am Main, pp. 29–35.

Asu, E. L. and Nolan, F. (2007), "Low accentuation in Estonian", *Language and Speech*, 50(4), pp.567–88.

Atterer, M. and Ladd, D. (2004), "On the phonetics and phonology of 'segmental anchoring' of f0: Evidence from German", *Journal of Phonetics*, 32, pp.177–97.

Aufterbeck, M. (2003), "Alignment of the trailing L tone in Scottish H*+L nuclei", *in Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, pp. 2937–40.

Baayen, R. (2008), *Analyzing linguistic data: A practical introduction to statistics using R*, CUP, Cambridge.

Baayen, R., Davidson, D. and Bates, D. (2008), "Mixed-effects modeling with crossed random effects for subjects and items", *Journal of Memory and Language*, 59, pp.390–412.

Bailey, G., Wikle, T., Tillery, J. and Sand, L. (1991), "The apparent time construct", *Language Variation and Change*, 3(3), pp.241–64.

Bakker, D., Müller, A., Velupillai, V., Wichmann, S., Brown, C. H., Brown, P., Egorov, D., Mailhammer, R., Grant, A. and Holman, E. W. (2009), "Adding typology to lexicostatistics: A combined approach to language classification", *Linguistic Typology*, 13, pp.169–81.

Bandelt, H.-J., Forster, P. and Röhl, A. (1999), "Median-joining networks for inferring intraspecific phylogenies", *Molecular Biology and Evolution*, 16, pp.37–48.

Bates, D. (2005), "Fitting linear mixed models in R", *R News*, 5(1), pp.27–31.

Baxter, W. H. and Manaster-Ramer, A. (2000), "Beyond lumping and splitting: Probabilistic issues in historical linguistics", *in* C. Renfrew, A. McMahon and L. Trask, eds, *Time depth in historical linguistics*, Vol. 1, The McDonald Institute for Archaeological Research, Cambridge, pp. 167–88.

Beckman, M. and Ayers Elam, G. (1997), 'Guidelines for ToBI labeling (version 3, march 1997)'. **URL:** *http://www.ling.ohio-state.edu/~tobi/ (Feb 2011)*

Beijering, K., Gooskens, C. and Heeringa, W. (2008), Phonetic and lexical predictors of intelligibility. Paper presented at the Thirteenth International Conference on Methods in Dialectology, Leeds.

Blevins, J. (2004), *Evolutionary phonology: The emergence of sound patterns*, CUP, Cambridge.

Blevins, J. (2007), "Interpreting misperception: Beauty is in the ear of the beholder", *in* M. J. Sole, P. Beddor and M. Ohala, eds, *Experimental approaches to phonology*, OUP, Oxford, pp. 144–54.

Blevins, J. and Garrett, A. (1998), "The origins of consonant-vowel metathesis", *Language*, 74(3), pp.508–56.

Boberg, C. (2009), "The emergence of a new phoneme: Foreign (a) in Canadian English", *Language Variation and Change*, 21(3), pp.355–80.

Boersma, P. and Weenink, D. (2009), 'Praat: doing phonetics by computer (version 5.0.47)[computer program]'. **URL:** *http://www.praat.org (Feb 2011)*

Bolinger, D. (1978), "Intonation across languages", *in* J. Greenberg, ed., *Universals of human language*, Vol. II: Phonology, Stanford University Press, Palo Alto, pp. 471–524.

Bolinger, D. (1989), *Intonation and its uses*, Stanford University Press, Palo Alto, CA.

Bradley, D. and Bradley, R. (1983), "Applications of sequence comparison to the study of bird songs", *in* D. Sankoff and J. B. Kruskal, eds, *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*, Addison-Wesley, Reading, MA, pp. 189–209.

Braun, B. (2006), "Phonetics and phonology of thematic contrast in German", *Language and Speech*, 49, pp.451–93.

Browman, C. P. and Goldstein, L. (1989), "Articulatory gestures as phonological units", *Phonology*, 49, pp.155–80.

Browman, C. P. and Goldstein, L. (1992), "Articulatory phonology: An overview", *Phonetica*, 6, pp.201–251.

Brown, G., Currie, K. L. and Kenworthy, J. (1980), *Questions of intonation*, Croom Helm, London.

Bruce, G. (1977), *Swedish word accents in sentence perspective*, CWK Gleerup, Lund.

Bruce, G. (1990), "Alignment and composition of tonal accents: Comments on Silverman and Pierrehumbert's paper", *in* J. Kingston and M. Beckman, eds, *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, CUP, Cambridge, pp. 72–106.

Bruce, G. (2007), "Components of a prosodic typology of Swedish intonation", *in* T. Riad and C. Gussenhoven, eds, *Tones and tunes. Vol. 1: Typological studies in word and sentence prosody*, Mouton de Gruyter, Berlin, pp. 113–46.

Bruce, G. and Gårding, E. (1978), "A prosodic typology for Swedish dialects", *in* E. Garding, G. Bruce and R. Bannert, eds, *Nordic prosody*, Gleerup, Lund, pp. 219–28.

Bryant, D. and Moulton, V. (2002), "Neighbornet: An agglomerative method for the construction of planar phylogenetic networks", *in* R. Guigó and D. Gusfield, eds, *Algorithms in Bioinformatics*, Springer, Berlin and Heidelberg, pp. 375–91.

Bye, P. (2004), Evolutionary typology and Scandinavian pitch accent. Unpublished article. **URL:** *http://www.hum.uit.no/a/bye/ (Feb 2011)*

Bye, P. (2008), The origin of double-peaked accent in North Germanic. Power-Point presentation for the Phonological Variation and Change FORSE PhD course, Flåm. **URL:** *http://www.hum.uit.no/a/bye/ (Feb 2011)*

Campbell, L. (2003), "How to show languages are related: Methods for distant genetic relationships", *in* B. D. Joseph and R. D. Janda, eds, *The Handbook of Historical Linguistics*, Blackwell, Malden, MA/Oxford/Carlton, Victoria, Australia, pp. 262–82.

Campbell, N. and Beckman, M. (1997), "Stress, prominence, and spectral tilt", *in Intonation: Theory, Models and Applications (Proceedings of an ESCA Workshop)*, Athens, pp. 67–70.

Caspers, J. and van Heuven, V. (1993), "Effects of time pressure on the phonetic realization of Dutch Accent-Lending pitch rise and fall", *Phonetica*, 50, pp.161–71.

Chambers, J. and Trudgill, P. (1980), *Dialectology*, CUP, Cambridge.

Chen, M. Y. (2000), *Tone Sandhi*, CUP, Cambridge.

Chen, Y. and Gussenhoven, C. (2008), "Emphasis and tonal implementation in Standard Chinese", *Journal of Phonetics*, 36, pp.724–46.

Clark, R. A. and Dusterhoff, K. E. (1999), "Objective methods for evaluating synthetic intonation", *in Proceedings of Eurospeech 1999*, Vol. 4, pp. 1623–6.

Clopper, C. G. and Paolillo, J. C. (2006), "North American English vowels: A Factor-analytic perspective", *Literary and Linguistic Computing*, 21(4), pp.445–62.

Colantoni, L. and Gurlekian, J. (2004), "Convergence and intonation: Historical evidence from Buenos Aires Spanish", *Bilingualism. Language and Cognition*, 7, pp.107–119.

Coleman, J. (2005), *Introducing speech and language processing*, CUP, Cambridge.

Connell, B. and Ladd, D. R. (1990), "Aspects of pitch realisation in Yoruba", *Phonology*, 7, pp.1–30.

Connolly, J. H. (1997), "Quantifying target-realization differences", *Clinical Linguistics and Phonetics*, 11, pp.267–98.

Cowlishaw, G. and Mace, R. (1996), "Cross-cultural patterns of marriage and inheritance: A phylogenetic approach", *Ethology and Sociobiology*, 17, pp.87–97.

Cruttenden, A. (1995), "Rises in English", *in* J. Windsor-Lewis, ed., *Studies in General and English phonetics: Essays in honour of Professor J.D. O'Connor*, Routledge, London, pp. 155–73.

Cruttenden, A. (1997), *Intonation*, 2nd edn, CUP, Cambridge.

Cruttenden, A. (2007), "Intonational diglossia: A case study of Glasgow", *Journal of the International Phonetic Association*, 37(2), pp.257–74.

Crystal, D. (1969), *Prosodic systems and intonation in English*, CUP, Cambridge.

Dalton, M. (2007), The phonetics and phonology of the intonation of Irish dialects, PhD thesis, Trinity College, Dublin.

Dalton, M. and Ní Chasaide, A. (2005), "Tonal alignment in Irish dialects", *Language and Speech*, 48(4), pp.441–64.

Dalton, M. and Ní Chasaide, A. (2007), "Melodic alignment and micro-dialect variation in Connemara Irish", *in* T. Riad and C. Gussenhoven, eds, *Tones and tunes. Vol. 2: Experimental studies in word and sentence prosody*, Mouton de Gruyter, Berlin, pp. 293–316.

D'arcy, S., Russell, M., Browning, S. and Tomlinson, M. (2004), "The Accents of the British Isles (ABI) corpus", *in Proceedings of Modélisations pour l'Identification des Langues*, Paris, pp. 115–9.

de Boer, E. (2009), What the split of the Tokyo type tone system into a number of sub-types can tell us about the tone system of proto-Japanese. Paper presented at the Nineteenth International Conference on Historical Linguistics (ICHL), Nijmegen.

de Jong, G., McDougall, K., Hudson, T. and Nolan, F. (2007), "The speaker discriminating power of sounds undergoing historical change: A formant-based study", *in Proceedings of the International Congress of Phonetic Sciences (ICPhS) XVI*, Saarbrücken, pp. 1813–6.

de Pijper, J. R. and Sanderman, A. A. (1994), "On the perceptual strength of prosodic boundaries and its relation to suprasegmental cues", *Journal of the Acoustical Society of America*, 96(4), pp.2037–47.

Delattre, P. (1965), *Comparing the phonetic features of English, French, German, and Spanish: An interim report*, Groot, Heidelberg.

Dilley, L. C. (2005), The phonetics and phonology of tonal systems, PhD thesis, MIT, Cambridge, MA.

Dilley, L. C., Ladd, D. and Schepman, A. (2005), "Alignment of L and H in bitonal pitch accents: Testing two hypotheses", *Journal of Phonetics*, 33, pp.115–9.

D'Imperio, M. (2000), The role of perception in defining tonal targets and their alignment, PhD thesis, The Ohio State University, Columbus, OH.

D'Imperio, M., Espresser, R., Loevenbruck, H., Menezes, C., Nguyen, N. and Welby, P. (2007), "Are tones aligned to articulatory events? Evidence from Italian and French", *in* J. I. Hualde and J. Cole, eds, *Papers in laboratory phonology IX*, Mouton de Gruyter, Berlin, pp. 577–607.

Dombrowski, E. and Niebuhr, O. (2005), "Acoustic patterns and communicative functions of phrase-final f0 rises in German: Activating and restricting contours", *Phonetica*, 62, pp.176–95.

Dunn, M., Terrill, A., Reesink, G., Foley, R. A. and Levinson, S. C. (2005), "Structural phylogenetics and the reconstruction of ancient language history", *Science*, 309, pp.2072–5.

Ellison, T. M. and Kirby, S. (2006), "Measuring language divergence by intralexical comparison", *in Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the ACL*, pp. 273–80.

Elordieta, G. and Calleja, N. (2005), "Microvariation in accentual alignment in Basque Spanish", *Language and Speech*, 48(4), pp.397–439.

Embleton, S. (1986), *Statistics in historical linguistics*, Brockmeyer, Bochum.

Embleton, S. (2000), "Lexicostatistics/glottochronology to Sankoff to Starostin to future horizons", *in* C. Renfrew, A. McMahon and L. Trask, eds, *Time depth in historical linguistics*, Vol. 1, The McDonald Institute for Archaeological Research, Cambridge, pp. 143–65.

Engstrand, O. and Nyström, G. (2002), "Meyer's accent contours revisited", *in Proceedings from Fonetik 2002, the 15th Swedish Phonetics Conference*, Vol. 44, pp. 17–20.

Fant, G. (1960), *Acoustic theory of speech production: With calculations based on X-ray studies of Russian articulations*, Mouton, s'Gravenhage.

Fant, G. (2004), *Speech acoustics and phonetics*, Kluwer Academic, Dordrecht/London.

Fant, G. and Kruckenberg, A. (2004), "Intonation analysis and synthesis with reference to Swedish", *in Proceedings of the International Symposium on Tonal Aspects of Language, TAL 2004*, Beijing, pp. 57–60.

Ferragne, E. and Pellegrino, F. (2010), "Vowel systems and accent similarity in the British Isles: Exploiting multidimensional acoustic distances in phonetics", *Journal of Phonetics*, 38(4), pp.526–39.

Fletcher, J., Grabe, E. and Warren, P. (2005), "Intonational variation in four dialects of English: The high rising tone", *in* S.-A. Jun, ed., *Prosodic typology: The phonology of intonation and phrasing*, OUP, Oxford, pp. 390–409.

Frisch, S. (1996), Similarity and frequency in phonology, PhD thesis, Northwestern University, Evanston, Illinois.

Glasberg, B. R. and Moore, B. C. (1990), "Derivation of auditory filter shapes from notched-noise data", *Hearing Research*, 47, pp.103–38.

Gleason, H. (1955), *An introduction to descriptive linguistics*, Henry Holt and company, New York.

Goebl, H. (1984), *Dialektometrische Studien anhand italoromanischer, rätoromanischer und galloromanischer Sprachmaterialien aus AIS und ALF*, M. Niemeyer, Tübingen.

Gooskens, C. (2005), "How well can Norwegians identify their dialects?", *Nordic Journal of Linguistics*, 28(1), pp.37–60.

Gooskens, C., Heeringa, W. and Beijering, K. (2008), "Phonetic and lexical predictors of intelligibility", *International Journal of Humanities and Arts Computing*, 2(1-2), pp.63–81.

Grabe, E. (1998), Comparative intonational phonology:English and German, PhD thesis, Max Planck Institute for Psycholinguistics, Nijmegen.

Grabe, E. (2002), "Variation adds to prosodic typology", *in Proceedings of the Speech Prosody Conference*, Aix-en-Provence, pp. 127–32.

Grabe, E., Coleman, J. and Kochanski, G. (2008), OXIGEN: The OXford Intonation GENerator. **URL:** *http://www.phon.ox.ac.uk/files/apps/oxigen/ (Feb 2011)*

Grabe, E., Kochanski, G. and Coleman, J. (2005), "The intonation of native accent varieties in the British Isles: Potential for miscommunication", *in* K. Dziubalska-Kolaczyk and J. Przedlacka, eds, *English pronunciation models: A changing scene*, Peter Lang, Bern, pp. 311–37.

Grabe, E., Kochanski, G. and Coleman, J. (2007), "Connecting intonation labels to mathematical descriptions of dundamental frequency", *Language and Speech*, 50(3), pp.281–310.

Grabe, E. and Post, B. (2002), "Intonational variation in the British Isles", *in Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence, pp. 343–6.

Grabe, E., Post, B. and Nolan, F. (2001), The IViE Corpus. Department of Linguistics, University of Cambridge. **URL:** *http://www.phon.ox.ac.uk/IViE/ (Feb 2011)*

Grabe, E., Post, B., Nolan, F. and Farrar, K. (2000), "Pitch accent realization in four varieties of British English", *Journal of Phonetics*, 28, pp.161–85.

Gray, R. D. and Atkinson, Q. D. (2003), "Language-tree divergence times support the Anatolian theory of Indo-European origin", *Nature*, 426, pp.435–9.

Gray, R. D., Drummond, A. and Greenhill, S. J. (2009), "Language phylogenies reveal expansion pulses and pauses in Pacific settlement", *Science*, 323, pp.429–83.

Gray, R. D., Greenhill, S. J. and Ross, R. M. (2007), "The pleasures and perils of darwinizing culture (with phylogenetics)", *Biological Theory*, 2(4), pp.360–75.

Gray, R. D. and Jordan, F. M. (2000), "Language trees support the express-train sequence of Austronesian expansion", *Nature*, 405, pp.1052–5.

Greenhill, S. J., Blust, R. and Gray, R. D. (2008), "The Austronesian Basic Vocabulary Database: From Bioinformatics to Lexomics", *Evolutionary Bioinformatics*, 4, pp.271–83.

Grice, M. (1995), "Leading tones and downstep in English", *Phonology*, 12(2), pp.183–233.

Grice, M., Ladd, D. and Arvaniti, A. (2000), "On the place of phrase accents in intonational phonology", *Phonology*, 17(2), pp.143–85.

Grimes, J. E. and Agard, F. B. (1959), "Linguistic divergence in Romance", *Language*, 35, pp.598–604.

Guion, S. G., Post, M. W. and Payne, D. L. (2004), "Phonetic correlates of tongue root vowel contrasts in Maa", *Journal of Phonetics*, 16, pp.355–69.

Gussenhoven, C. (1984), *On the grammar and semantics of sentence accents*, Foris, Dordrecht, The Netherlands.

Gussenhoven, C. (2004), *The phonology of tone and intonation*, CUP, Cambridge.

Gussenhoven, C. (2007), "The phonology of intonation", *in* P. de Lacy, ed., *The Handbook of Phonology*, CUP, Cambridge, pp. 253–80.

Gussenhoven, C. and Rietveld, A. (1991), "An experimental evaluation of two nuclear-tone taxonomies", *Linguistics*, 29, pp.423–449.

Gussenhoven, C. and Rietveld, T. (2000), "The behavior of H* and L* under variations in prominence of $f0$ peaks", *Language and Speech*, 43, pp.183–203.

Haan, J. (2002), *Speaking of questions*, LOT, Utrecht.

Halliday, M. (1967), *Intonation and grammar in British English*, Mouton, The Hague.

Hardcastle, W. J. and Laver, J., eds (1997), *The handbook of phonetic sciences*, Blackwell, Oxford/Cambridge, MA.

Hardcastle, W. J., Laver, J. and Gibbon, F. E., eds (2010), *The handbook of phonetic sciences*, 2nd edn, Blackwell/Wiley, Chichester, West Sussex.

Hayes, B. and Lahiri, A. (1991), "Bengali intonational phonology", *Natural Language and Linguistic Theory*, 9(1), pp.47–96.

Hayward, R. A. (2000), "Is there a metric for convergence?", *in* C. Renfrew, A. McMahon and L. Trask, eds, *Time depth in historical linguistics*, Vol. 2, The McDonald Institute for Archaeological Research, Cambridge, pp. 621–40.

Heeringa, W. (2004), Measuring dialect pronunciation difference using Levenshtein distance, PhD thesis, University of Groningen, Groningen.

Heeringa, W. and Braun, A. (2003), "The use of the Almeida-Braun system in the measurements of Dutch dialect differences", *Computers and the Humanities*, 37(3), pp.257–71.

Heeringa, W., Johnson, K. and Gooskens, C. (2009), "Measuring Norwegian dialect distances using acoustic features", *Speech Communication*, 51(2), pp.167–83.

Heeringa, W. and Joseph, B. (2007), "The relative divergence of Dutch dialect pronunciations from their common source: An exploratory study", *in Proceedings of the Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*, Prague, pp. 31–9.

Heeringa, W. and Nerbonne, J. (2001), "Dialect areas and dialect continua", *Language variation and change*, 13, pp.375–400.

Heggarty, P. (2000a), Quantification and comparison in language structure, PhD thesis, University of Cambridge.

Heggarty, P. (2000b), "Quantifying change over time in phonetics", *in* C. Renfrew, A. McMahon and L. Trask, eds, *Time depth in historical linguistics*, Vol. 2, The McDonald Institute for Archaeological Research, Cambridge, pp. 531–62.

Heggarty, P. (2006), "Interdisciplinary indiscipline? Can phylogenetic methods be meaningfully applied to language data-and to dating language?", *in* P. Forster and C. Renfrew, eds, *Phylogenetic methods and the prehistory of languages*, McDonald Institute for Archaeological Research, Cambridge, pp. 183–94.

Heggarty, P. (2010), "Beyond lexicostatistics: How to get more out of 'word list' comparisons", *Diachronica*, 27(2), pp.301–32.

Heggarty, P., Maguire, W. and McMahon, A. (2010), "Splits or waves? Trees or webs? How divergence measures and network analysis can unravel language histories", *Philosophical Transactions of the Royal Society (B)*, 365, pp.3829–43. Special edition on Cultural and linguistic Diversity: Evolutionary approaches.

Heggarty, P., McMahon, A. and McMahon, R. (2005), "From phonetic similarity to dialect classification: A principled approach", *in* N. Delbecque, J. van der Auwera and D. Geeraerts, eds, *Persepctives on variation: Sociolinguistic, historical, comparative*, Mouton de Gruyter, Berlin, pp. 43–91.

Herman, R. and McGory, J. (2002), "The conceptual similarity of intonational tones and its effects on intertranscriber reliability", *Language and Speech*, 45(1), pp.1–36.

Hermes, D. J. (1998), "Auditory and visual similarity of pitch contours", *Journal of Speech, Language, and Hearing Research*, 41, pp.63–72.

Hermes, D. J. and van Gestel, J. C. (1991), "The frequency scale of speech intonation", *Journal of the Acoustical Society of America*, 90(1), pp.97–102.

Hirst, D. and di Cristo, A. (1998), *Intonation systems: A survey of twenty languages*, CUP, Cambridge.

Hognestad, J. K. (2006), "Tonal accents in Stavanger: From Western towards Eastern Norwegian prosody?", *in* G. Bruce and M. Horne, eds, *Nordic Prosody: Proceedings of the IXth Conference, Lund*, Peter Lang, Frankfurt am Main, pp. 107–116.

Holder, M. and Lewis, P. O. (2003), "Phylogeny estimation: Traditional and Bayesian approaches", *Genetics*, 4, pp.275–84.

Holmes, J. and Holmes, W. (1995), *Speech synthesis and recognition*, 2nd edn edn, Taylor & Francis, London/New York.

Horne, M., ed. (2000), *Prosody, theory, and experiment: Studies presented to Gösta Bruce*, Kluwer Academic Publishers, Dordrecht/London.

House, D. (2003), "Perceiving question intonation: The role of pre-focal pause and delayed focal peak", *in Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, pp. 755–8.

House, J. and Wichmann, A. (1996), "Investigating peak timing in naturally-occurring speech: From segmental constraints to discourse structure", *Speech, Hearing and Language: Work in Progress*, 9, pp.99–117.

House, J., Wichmann, A. and Rietveld, T. (2000), "Discourse constraints on $f0$ peak timing in English", *in* A. Botinis, ed., *Intonation: Analysis, modelling and technology*, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 163–82.

Hualde, J. I. (2007), "Historical convergence and divergence in Basque accentuation", *in* T. Riad and C. Gussenhoven, eds, *Tones and tunes. Vol. 1: Typological studies in word and sentence prosody*, Mouton de Gruyter, Berlin, pp. 291–32.

Huckvale, M. (2004), "ACCDIST: A metric for comparing speakers' accents", *in Proceedings of the International Conference on Spoken Language Processing*, Jeju, Korea. **URL:** *http://www.phon.ucl.ac.uk/home/mark/ (Feb 2011)*

Huckvale, M. (2007), "Hierarchical clustering of speakers into accents with the ACCDIST metric", *in Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, Saarbrüken. **URL:** *http://www.phon.ucl.ac.uk/home/mark/ (Feb 2011)*

Hughes, E. (1956), *North Country life in the 18th Century.*, Vol. 1: The North-East 1700-1750, OUP, London.

Huson, D. and Bryant, D. (2006), "Application of phylogenetic networks in evolutionary studies", *Molecular Biology and Evolution*, 23(2), pp.254–67.

Hyman, L. M. (2007), "Universals of tone rules: 30 years later", *in* T. Riad and C. Gussenhoven, eds, *Tones and tunes. Vol. 1: Typological studies in word and sentence prosody*, Mouton de Gruyter, Berlin, pp. 1–34.

Jakobson, R., Gunnar, C., Fant, M. and Halle, M. (1952), *Preliminaries to speech analysis: The distinctive features and their correlates*, Acoustics Lab, MIT, Cambridge, MA.

Jarman, E. and Cruttenden, A. (1976), "Belfast intonation and the myth of the fall", *Journal of the International Phonetic Association*, 6, pp.4–12.

Joseph, B. D. and Janda, R. D., eds (2003), *Handbook of historical linguistics*, Blackwell, Malden, MA/Oxford/Carlton, Victoria, Australia.

Jurafsky, D. and Martin, J. H. (2008), *Speech and language processing: An introduction to Natural Language Processing*, 2nd edn, Prentice Hall, Upper Saddle River, NJ.

Kalaldeh, R., Dorn, A. and Ní Chasaide, A. (2009), "Tonal alignment in three varieties of Hiberno-English", *in Proceedings of Interspeech*, Brighton, pp. 2443–6.

Kerswill, P. (2003), "Dialect levelling and geographical diffusion in British English", *in* D. Britain and J. Cheshire, eds, *Social dialectology. In honour of Peter Trudgill*, Benjamins, Amsterdam, pp. 223–43.

Kessler, B. (1995), "Computational dialectology in Irish Gaelic", *in Proceedings of the 7th Conference of the European Chapter of the Association for Computational Linguistics*, Dublin, pp. 60–7.

Kessler, B. (2001), *The significance of word lists*, CSLI Publications, Stanford.

Kessler, B. (2005), "Phonetic comparison algorithms", *Transactions of the Philological Society*, 103(2), pp.243–60.

Kessler, B. (2007), "Word similarity metrics and multilateral comparison", *in Proceedings of the Ninth Meeting of the ACL Special Interest Group in Computational Morphology and Phonology*, Prague, pp. 6–14.

Kessler, B. and Lehtonen, A. (2006), "Multilateral comparison and significance testing of the Indo-Uralic question", *in* P. Forster and C. Renfrew, eds, *Phylogenetic methods and the prehistory of languages*, The McDonald Institute for Archaeological Research, Cambridge, pp. 33–42.

Knight, R.-A. and Nolan, F. (2006), "The effect of pitch span on intonational plateaux", *Journal of the International Phonetic Association*, 36(1), pp.21–38.

Knowles, G. (1975), Scouse: the urban dialect of Liverpool, PhD thesis, University of Leeds.

Knowles, G. (1978), "The nature of phonological variables in Scouse", *in* P. Trudgill, ed., *Sociolinguistic patterns in British English*, Edward Arnold, London, pp. 80–90.

Knowles, G. (1984), "Variable strategies in intonation", *in* D. Gibbon and H. Richter, eds, *Intonation, accent and rhythm: Studies in Discourse Phonology*, Walter de Gruyter, Berlin/New York, pp. 226–42.

Kochanski, G., Grabe, E., Coleman, J. and Rosner, B. (2005), "Loudness predicts prominence: Fundamental frequency lends little", *Journal of the Acoustical Society of America*, 118(2), pp.1038–54.

Kochanski, G. and Shih, C. (2003), "Prosody modeling with soft templates", *Speech Communication*, 39(3-4), pp.311–52.

Kondrak, G. (2000), "A new algorithm for the alignment of phonetic sequences", *in Proceedings of the First Meeting of the North American Chapter of the Association for Computational Linguistics*, Seattle, pp. 288–95.

Kondrak, G. (2005), "N-gram similarity and distance.", *in Proceedings of the Twelfth International Conference on String Processing and Information Retrieval (SPIRE 2005)*, Buenos Aires, pp. 115–126.

Kondrak, G. and Sherif, T. (2006), "Evaluation of several phonetic similarity algorithms on the task of cognate identification", *in Proceedings of the COLING-ACL Workshop on Linguistic Distances*, Sydney, pp. 43–50.

Kristoffersen, G. (2000), *The phonology of Norwegian*, OUP, Oxford.

Kristoffersen, G. (2006), "Tonal melodies and tonal alignment in East Norwegian", *in* G. Bruce and M. Horne, eds, *Nordic Prosody: Proceedings of the IXth Conference, Lund*, Peter Lang, Frankfurt am Main, pp. 157–66.

Kristoffersen, G. (2007), "Dialect variation in East Norwegian tone", *in* T. Riad and C. Gussenhoven, eds, *Tones and tunes. Vol 1: Typological studies in word and sentence prosody*, Mouton de Gruyter, Berlin/New York, pp. 99–111.

Kroch, A. (2001), "Syntactic change", *in* M. Baltin and C. Collins, eds, *The handbook of contemporary syntactic theory*, Blackwell, Oxford, pp. 699–729.

Kroeber, A. and Chrétien, C. (1937), "Quantitative classification of Indo-European languages", *Language*, 13(2), pp.83–103.

Kruskal, J. B. (1983), "An overview of sequence comparison", *in* D. Sankoff and J. B. Kruskal, eds, *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*, Addison-Wesley, Reading, MA, pp. 1–44.

Kruskal, J. B. and Liberman, M. (1983), "The symmetric time-warping problem; from continuous to discrete", *in* D. Sankoff and J. B. Kruskal, eds, *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*, Addison-Wesley, Reading, MA, pp. 125–59.

Labov, W. (1963), "The social motivation of a sound change", *Word*, 19(3), pp.273–309.

Labov, W. (1966), *The social stratification of English in New York City*, Center for Applied Linguistics, Washington D.C.

Labov, W. (1981), "Resolving the Neogrammarian controversy", *Language*, 57(2), pp.267–308.

Labov, W. (1994), *Principles of linguistic change*, Vol. I: Internal factors, Blackwell, Oxford.

Labov, W. (2007), "Transmission and diffusion", *Language*, 83(2), pp.344–87.

Labov, W., Ash, S. and Boberg, C. (2006), *The atlas of North American English*, Mouton de Gruyter, Berlin.

Ladd, D. R. (1983), "Phonological features of intonational peaks", *Language*, 59, pp.721–59.

Ladd, D. R. (2006), "Segmental anchoring of pitch movements: Autosegmental phonology or gestural coordination", *Rivista di Linguistica (Italian Journal of Linguistics)*, 18, pp.19–38.

Ladd, D. R. (2008), *Intonational phonology*, CUP, Cambridge. 2nd edn.

Ladd, D. R. (forthcoming), "Phonetics in phonology", *in* J. Goldsmith, J. Riggle and A. Yu, eds, *Handbook of phonological theory*, 2nd edn, Blackwell. **URL:** *www.ling.ed.ac.uk/~bob/ (Feb 2011)*

Ladd, D. R., Faulkner, D., Faulkner, H. and Schepman, A. (1999), "Constant "segmental anchoring" of *f*0 movements under changes in speech rate", *Journal of the Acoustical Society of America*, 106, pp.1543–54.

Ladd, D. R., Mennen, I. and Schepman, A. (2000), "Phonological conditioning of peak alignment in rising pitch accents in Dutch", *Journal of the Acoustical Society of America*, 107(5), pp.2685–96.

Ladd, D. R. and Morton, R. (1997), "The perception of intonational emphasis: Continuous or categorical", *Journal of Phonetics*, 25, pp.313–42.

Ladd, D. R. and Schepman, A. (2003), ""Sagging transitions" between high pitch accents in English: experimental evidence", *Journal of Phonetics*, 31, pp.81–112.

Ladd, D. R., Schepman, A., White, L., Quarmby, L. M. and Stackhouse, R. (2009), "Structural and dialectal effects on pitch peak alignment in two varieties of British English", *Journal of Phonetics*, 37(2), pp.145–61.

Ladefoged, P. (1975), *A course in phonetics*, Harcourt Brace Jovanovich, New York.

Ladefoged, P. (2006), *A course in phonetics*, 5th edn, Thomson Wadsworth, Boston.

Ladefoged, P. and Maddieson, I. (1996), *The sounds of the world's languages*, Blackwell, Oxford & Cambridge, MA.

Laver, J. (1994), *Principles of phonetics*, CUP, Cambridge.

Levenshtein, V. (1966), "Binary codes capable of correcting deletions, insertions and reversals", *Cybernetics and Control Theory*, 10(8), pp.707–10.

Liberman, M. (1979), *The intonational system of English*, Garland Pub. Inc., New York & London. Published version of 1975 MIT Doctoral thesis.

Liberman, M. and Pierrehumbert, J. (1984), "Intonational invariance under changes in pitch range and length", *in* M. Aronoff and R. T. Oehrle, eds, *Language Sound Structure*, The MIT Press, Cambridge, MA & London, pp. 157–233.

Lickley, R. J., Schepman, A. and Ladd, D. R. (2005), "Alignment of "phrase accent" lows in Dutch falling rising questions: Theoretical and methodological implications", *Language and Speech*, 48(2), pp.157–83.

Liljencrants, J. and Lindblom, B. (1972), "Numerical simulation of vowel quality systems: The role of perceptual contrast", *Language*, 48(4), pp.839–62.

Lindblom, B. and Maddieson, I. (1988), "Phonetic universals in consonant systems", *in* L. M. Hyman and C. N. Li, eds, *Language, speech, and mind*, Routledge, London, pp. 62–78.

Lohr, M. (2000), "New approaches to lexicostatistics and glottochronology", *in* C. Renfrew, A. McMahon and L. Trask, eds, *Time depth in historical linguistics*, Vol. 2, The McDonald Institute for Archaeological Research, Cambridge, pp. 209–22.

Longobardi, G. and Guardiano, C. (2009), "Evidence for syntax as a signal of historical relatedness", *Lingua*, 119, pp.1679–1706.

Lowry, O. (1997), Intonation patterns in Northern Irish English. MPhil thesis, University of Cambridge.

Lowry, O. (2002*a*), Belfast intonation: Testing the ToBI framework of intonational analysis, PhD thesis, University of Ulster, Newtownabbey, Northern Ireland.

Lowry, O. (2002*b*), "The stylistic variation of nuclear patterns in Belfast English", *Journal of the International Phonetic Association*, 32(1), pp.33–42.

Maddieson, I. (1984), *Patterns of sounds*, CUP, Cambridge.

Maguire, W. (2008), "Quantifying dialect similarity by comparison of the lexical distribution of phonemes", *International Journal of Humanities and Arts Computing*, 2(1-2), pp.261–277.

Maguire, W., McMahon, A., Heggarty, P. and Dediu, D. (2010), "The past, present, and future of English dialects: Quantifying convergence, divergence, and dynamic equilibrium", *Language Variation and Change*, 22, pp.69–104.

Makarova, V. (2007), "The effect of pitch peak alignment on sentence type identification in Russian", *Language and Speech*, 50(3), pp.385–422.

Mayo, C. (1996), Prosodic transcription of Glasgow English: An evaluation study, Master's thesis, University of Edinburgh.

Mayo, C., Aylett, M. and Ladd, D. R. (1997), "Prosodic transcription of Glasgow English: An evaluation study", *in Intonation: Theory, Models, and Applications*, pp. 231–4.

IPA (1999), *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*, CUP, Cambridge.

McElholm, D. D. (1986), "Intonation in Derry English: A preliminary study", *in* H. Kirkwood, ed., *Studies in intonation (Occasional Papers in Linguistics and Language Learning No.11 June 1986*, University of Ulster, Ulster, pp. 1–58.

McGregor, J. G. (2005), High rising tunes in Australian English, PhD thesis, Macquarie University, Sydney.

McLemore, C. A. (1991), The pragmatic interpretation of English intonation: Sorority speech, PhD thesis, University of Texas, Austin.

McMahon, A. (1991), "Lexical phonology and sound change: The case of the Scottish Vowel Length Rule", *Journal of Linguistics*, 27(1), pp.29–53.

McMahon, A. (2004), "Prosodic change and language contact", *Bilingualiam: Language and cognition*, 7(2), pp.121–3. Review of Colantoni and Gurlekian (2004-this volume).

McMahon, A. (2007*a*), "Sounds, brain and evolution: Or, why phonology is plural", *in* M. C. Pennington, ed., *Phonology in context*, Palgrave Macmillan, Basingstoke/New York, pp. 159–85.

McMahon, A. (2007*b*), "Who's afraid of the vowel shift rule?", *Language Sciences*, 29, pp.341–59.

McMahon, A. (2008), From the Bard to Bob Marley: Comparing the sounds of accents of English, past and present. Inaugural Lecture, University of Edinburgh (Feb 2008).

McMahon, A., Foulkes, P. and Tollfree, L. (1994), "Gestural representation and Lexical Phonology", *Phonology*, 11(2), pp.277–316.

McMahon, A., Heggarty, P., McMahon, R. and Maguire, W. (2007), "The sound patterns of Englishes: Representing phonetic similarity", *English Language and Linguistics*, 11(1), pp.113–42.

McMahon, A., Heggarty, P., McMahon, R. and Slaska, N. (2005), "Swadesh sublists and the benefits of borrowing: An Andean case study", *Transactions of the Philological Society*, 103(2), pp.147–70.

McMahon, A. M. (1994), *Understanding language change*, CUP, Cambridge/New York/Melbourne.

McMahon, A. M. (2000), *Lexical phonology and the history of English*, CUP, Cambridge.

McMahon, A., Maguire, W. and Heggarty, P. (2005-07), Sound comparisons: Dialect and language comparison and classification by phonetic similarity. **URL:** *http://www.soundcomparisons.com (Feb 2011)*

McMahon, A. and McMahon, R. (2005), *Language classification by numbers*, OUP, Oxford.

McMahon, A. and McMahon, R. (2006), "Keeping contact in the family: Approaches to language classification and contact-induced change", *in* Y. Matras, A. McMahon and N. Vincent, eds, *Linguistic Areas: Convergence in historical and typological perspective*, Palgrave Macmillan, Basingstoke, Hampshire, pp. 51–74.

Mielke, J. (2005), Modeling distinctive feature emergence. Presentation given at West Coast Conference on Formal Linguistics (WCCFL) XXIV. 281-9. **URL:** *http://aix1.uottawa.ca/˜jmielke/ (Feb 2011)*

Mielke, J. and Roy, J. (2008), Phonetic similarity in ultrasound: Comparing deterministic, statistical, and information-theoretic approaches. Presentation given at MOT. **URL:** *http://aix1.uottawa.ca/˜jmielke/ (Feb 2011)*

Miller, G. A. and Nicely, P. E. (1955), "An analysis of perceptual confusions among some English consonants", *Journal of the Acoustical Society of America*, 27(2), pp.338–52.

Mills, T. I. P. (2008), Speech motor control variables in the production of voicing contrasts and emphatic accent, PhD thesis, University of Edinburgh.

Moisl, H. and Maguire, W. (2008), "Identifying the main determinants of phonetic variation in the Newcastle Electronic Corpus of Tyneside English", *Journal of Quantitative Linguistics*, 15(1), pp.46–69.

Morton, J. (2006), *Sound Recording: The life story of a technology*, Johns Hopkins University Press, Baltimore.

Mücke, D., Grice, M., Becker, J. and Hermes, A. (2009), "Sources of variation in tonal alignment: Evidence from acoustic and kinematic data", *Journal of Phonetics*, 37(3), pp.321–38.

Mücke, D. and Hermes, A. (2007), "Phrase boundaries and peak alignment: An acoustic and articulatory study", *in Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, pp. 997–1000.

Mücke, D., Nam, H., Hermes, A. and Goldstein, L. (to appear), "Coupling of tone and constriction gestures in pitch accents", *in Consonant clusters and structural complexity*. Draft. **URL:** *http://www.uni-koeln.de/phil-fak/phonetik/Institut/Mitarbeiter/dmuecke/ (Feb 2011)*

Nakhleh, L., Ringe, D. and Warnow, T. (2005), "Perfect phylogenetic networks: A new method for reconstructing the evolutionary history of natural languages", *Language*, 81, pp.382–420.

Nakhleh, L., Warnow, T., Ringe, D. and Evans, S. N. (2005), "A comparison of phylogenetics reconstruction methods on an Indo-European dataset", *Transactions of the Philological Society*, 103(2), pp.171–92.

Nerbonne, J. (2007), "Review of McMahon and McMahon (2005)", *Linguistic Typology*, 11, pp.425–36.

Nerbonne, J. and Heeringa, W. (1997), "Measuring dialect difference phonetically", *in* J. Coleman, ed., *Workshop on computational phonology*, Special Interest Group of the Association for Computational Linguistics, Madrid, pp. 11–18.

Nerbonne, J. and Heeringa, W. (2010), "Measuring dialect differences", *in* J. E. Schmidt and P. Auer, eds, *Theories and methods*, Mouton De Gruyter, Berlin, pp. 550–67.

Nerbonne, J., Heeringa, W. and Kleiwig, P. (1999), "Edit distance and dialect proximity", *in* D. Sankoff and J. B. Kruskal, eds, *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*, 1999 edn, CSLI Press, Stanford, pp. v–xv.

Ní Chasaide, A. and Dalton, M. (2006), "Dialect alignment signatures", *in Proceedings of Speech Prosody*, Dresden. **URL:** *http://www.tara.tcd.ie/jspui/handle/2262/39396 (Feb 2011)*

Nolan, F. (1997), "Speaker recognition and forensic phonetics", *in* W. Hardcastle and J. Laver, eds, *A Handbook of Phonetic Science*, Blackwell, Oxford, pp. 744–67.

Nolan, F. (2003), "Intonational equivalence: An experimental evaluation of pitch scales", *in Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, pp. 771–4.

Nolan, F. and Farrar, K. (1999), "Timing of f0 peaks and peak lag", *in Proceedings of the 14th International Congress of Phonetic Sciences (ICPhS)*, San Francisco, pp. 961–4.

O'Connor, J. D. and Arnold, G. F. (1961), *Intonation of colloquial English: A practical handbook*, Longman, London.

Oswalt, R. L. (1970), "The detection of remote linguistic relationships", *Computer Studies*, 3, pp.117–29.

Patel, A. D., Iverson, J. R. and Rosenberg, J. C. (2006), "Comparing the rhythm and melody of speech and music: The case of British English and French", *Journal of the Acoustical Society of America*, 119(5), pp.3034–47.

Patterson, D. (2000), A linguistic approach to pitch range modelling, PhD thesis, University of Edinburgh.

Pellowe, J. and Jones, V. (1978), "On intonational variability in Tyneside speech", *in* P. Trudgill, ed., *Sociolinguistic patterns in British English*, Edward Arnold, London, pp. 101–121.

Pellowe, J., Nixon, G., Strang, B. and McNeany, V. (1972), "A dynamic modelling of linguistic variation: The Urban (Tyneside) Linguistic Survey", *Lingua*, 30, pp.1–30.

Perkell, J., Cohen, M., Matthies, M., Garabieta, I. and Jackson, M. (1992), "Electro-magnetic midsagittal articulometer (EMMA) systems for transducing speech articulatory movements", *Journal of the Acoustical Society of America*, 92, pp.3078–96.

Peters, J. (2007), "Bitonal lexical pitch accents in the Limburgian dialect of Borgloon", *in* T. Riad and C. Gussenhoven, eds, *Tones and tunes. Vol. 1: Typological studies in word and sentence prosody*, Mouton de Gruyter, Berlin, pp. 167–98.

Peters, J., Gilles, P., Auer, P. and Selting, M. (2002), "Identification of regional varieties by intonational cues. an experimental study on Hamburg and Berlin German", *Language and Speech*, 45(2), pp.115–39.

Peters, J., Gilles, P., Auer, P. and Selting, M. (2003), "Identifying regional varieties by pitch information: A comparison of two approaches", *in Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, pp. 1065–8.

Piercy, C. (2010), One /a/ or two? The phonetics, phonology and sociolinguistics of change in the TRAP and BATH vowels in the southwest of England, PhD thesis, University of Essex.

Pierrehumbert, J. B. (1980), The phonetics and phonology of English intonation, PhD thesis, MIT.

Pierrehumbert, J. B. and Steele, S. A. (1989), "Categories of tonal alignment in English", *Phonetica*, 46(4), pp.181–96.

Pierrehumbert, J. and Beckman, M. (1988), *Japanese tone structure*, Cambridge, MA, The MIT Press.

Pike, K. L. (1945), *The intonation of American English*, University of Michigan Press, Ann Arbor, MI.

Pisoni, D. (1973), "Auditory and phonetic memory codes in the discrimination of consonants and vowels", *Perception & Psychophysics*, 13(2), pp.253–60.

Port, B. and Leary, A. (2005), "Against formal phonology", *Language*, 81(4), pp.927–64.

Prieto, P., D'Imperio, M. and Fivela, B. G. (2005), "Pitch accent alignment in Romance: Primary and secondary associations with metrical structure", *Language and Speech*, 48(4), pp.359–96.

Prieto, P., Mücke, D., Becker, J. and Grice, M. (2007), "Coordination patterns between pitch movements and oral gestures in Catalan", *in Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS)*, Saarbrücken, pp. 989–92.

Prieto, P. and Torreira, F. (2007), "The segmental anchoring hypothesis revisited: Syllable structure and speech rate effects on peak timing in Spanish", *Journal of Phonetics*, 35, pp.473–500.

Prieto, P., van Santen, J. and Hirschberg, J. (1995), "Tonal alignment patterns in Spanish", *Journal of Phonetics*, 23, pp.429–51.

Prokic, J. and Nerbonne, J. (2009), Phylogenetic inference based on sound correspondences. Paper presented at the 19th International Conference on Historical Linguistics, Nijmegen.

Queen, R. M. (2001), "Bilingual intonation patterns: Evidence of language change from Turkish-German bilingual children", *Language in Society*, 30, pp.55–80.

Quené, H. and van den Bergh, H. (2004), "On multi-level modeling of data from repeated measures designs: a tutorial", *Speech Communication*, 43, pp.103–121.

Quené, H. and van den Bergh, H. (2008), "Examples of mixed-effects modeling with crossed random effects and with binomial data", *Journal of Memory and Language*, 59, pp.413–425.

R Development Core Team (2010), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rahilly, J. (1991), Intonation patterns in normal hearing and postlingually deafened adults in Belfast, PhD thesis, Queen's University, Belfast.

Remijsen, B. and van Heuven, V. J. (1999), "Gradient and categorical pitch dimensions in Dutch: Diagnostic test", *in Proceedings of the International Congress of Phonetic Sciences (ICPhS)*, Vol. 3, San Francisco, pp. 1865–8.

Renfrew, C. and Forster, P. (2006), "Introduction", *in* P. Forster and C. Renfrew, eds, *Phylogenetic methods and the prehistory of languages*, McDonald Institute for Archaeological Research, Cambridge, pp. 1–9.

Repp, B. H. (1982), "Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception", *Psychological Bulletin*, 92(1), pp.81–110.

Riad, T. (1996), "Remarks on the Scandinavian tone accent typology", *Nordlyd*, 24, pp.129–56.

Rietveld, A. and Gussenhoven, C. (1985), "On the relation between pitch excursion size and prominence", *Journal of Phonetics*, 13, pp.299–308.

Rietveld, A. and Gussenhoven, C. (1995), "Aligning pitch targets in speech synthesis: Effects of syllable structure", *Journal of Phonetics*, 23, pp.375–85.

Ringe, D. (1992), "On calculating the factor of chance in language comparison", *Transactions of the American Philosophical Society*, 82(1).

Ringe, D., Warnow, T. and Taylor, A. (2002), "Indo-European and computational cladistics", *Transactions of the Philological Society*, 100(1), pp.59–129.

Rossi, M. (1971), "Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole", *Phonetica*, 23, pp.1–33.

Sankoff, D. and Kruskal, J. B., eds (1983), *Time warps, string edits, and macromolecules: The theory and practice of sequence comparison*, Addison-Wesley, Reading, MA.

Schepman, A., Lickley, R. and Ladd, D. R. (2006), "Effects of vowel length and "right context" on the alignment of Dutch nuclear accents", *Journal of Phonetics*, 34(1), pp.1–28.

Segerup, M. and Nolan, F. (2006), "Gothenburg Swedish word accents: A case of cue trading?", *in* G. Bruce and M. Horne, eds, *Nordic Prosody: Proceedings of the IXth Conference, Lund*, Peter Lang, Frankfurt am Main, pp. 225–33.

Séguy, J. (1973), "La dialectométrie dans l'Atlas linguistique de la Gascogne", *Revue de linguistique romane*, 37, pp.1–24.

Serva, M. and Petroni, F. (2008), "Indo-European languages tree by Levenshtein distance", *Europhysics Letters*, 81. Paper 68005.

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

Shobbrook, K. and House, J. (2003), "High rising tones in Southern British English", *in Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS)*, Barcelona, pp. 1273–6.

Shosted, R., Del-Giudice, A. and Arvaniti, A. (2006), Comparing methods for measuring pitch "elbows". Poster presented at the 4th Joint Meeting of the Acoustical Society of America (ASA) and the Acoustic Society of Japan, Honolulu.

Silverman, D. (1997), "Tone sandhi in Comaltepec Chinantec", *Language*, 73, pp.473–92.

Silverman, K., Beckman, M., Pitrelli, J., Ostendorf, M., Wightman, C., Price, P., Pierrehumbert, J. and Hirschberg, J. (1992), "ToBI: A standard for labelling English prosody", *in Proceedings of the 2nd International Conference on Spoken Language Processing (ICSLP)*, Banff, Alberta, pp. 867–70.

Silverman, K. and Pierrehumbert, J. (1990), "The timing of prenuclear high accents in English", *in* J. Kingston and M. Beckman, eds, *Papers in Laboratory Phonology I: Between the grammar and physics of speech*, CUP, Cambridge, pp. 72–106.

Slater, A. (2008), The IViE Corpus: English intonation in the British Isles. **URL:** *http://www.phon.ox.ac.uk/files/apps/IViE/index.php (Feb 2011)*

Somers, H. L. (1999), "Aligning phonetic segments for children's articulation assessment", *Computational Linguistics*, 25(2), pp.267–75.

Steele, S. A. (1986), "Nuclear accent *f*0 peak location: Effects of rate, vowel, and number of following syllables", *Journal of the Acoustical Society of America*, 80(Suppl. 1), pp.S51.

Stevens, K. N. (2002), "Toward a model for lexical access based on acoustic landmarks and distinctive features", *Journal of the Acoustical Society of America*, 111(4), pp.1872–91.

Stockwell, R. P. and Minkova, D. (2001), *English words: History and structure*, Cambridge, CUP.

Sudbury, A. (2001), "Falkland Islands English: A southern hemisphere variety?", *English World-Wide*, 22(1), pp.55–80.

Sullivan, J. and McMahon, A. (2010), "Phonetic comparison, varieties and networks: Swadesh's influence lives on here too", *Diachronica*, 27(2), pp.325–40.

Swadesh, M. (1950), "Salish internal relationships", *International Journal of American Linguistics*, 21, pp.121–37.

Swadesh, M. (1972), *The origin and diversification of language*, Routledge & Kegan Paul, London. ed. Joel Sherzer.

't Hart, J., Collier, R. and Cohen, A. (1990), *A perceptual study of intonation: An experimental-phonetic approach to speech melody*, CUP, Cambridge.

Tang, C. and van Heuven, V. (2007), "Mutual intelligibility and distance of chinese dialects. predicting judgments from objective measures", *in* B. Los and M. Koppen, eds, *Linguistics in the Netherlands 2007*, Benjamins, pp. 223–234.

Taylor, P. (1992), A phonetic model of English intonation, PhD thesis, University of Edinburgh.

Taylor, P. (2000), "Analysis and synthesis of intonation using the Tilt model", *Journal of the Acoustical Society of America*, 107(3), pp.1697–1714.

Torgersen, E. and Kerswill, P. (2004), "Internal and external motivation in phonetic change: Dialect levelling outcomes for an English vowel shift", *Journal of Sociolinguistics*, 8, pp.24–53.

Trager, G. and Smith, H. (1951), *An outline of English structure*, Battenburg Press, Norman, OK.

Turk, A., Nakai, S. and Sugahara, M. (2006), "Acoustic segment durations in prosodic research: A practical guide", *in* S. Sudhoff, D. Lenertová, R. Meyer, S. Pappert, P. Augurzky, I. Mleinek, N. Richter and J. S. r, eds, *Methods in empirical prosody research*, De Gruyter, Berlin/New York, pp. 1–28.

van Heuven, V. and van Zanten, E. (2005), "Speech rate as a secondary prosodic characteristic of polarity questions in three languages", *Speech Communication*, 47, pp.87–99.

van Leyden, K. and van Heuven, V. (2006), "On the prosody of Orkney and Shetland dialects", *Phonetica*, 63, pp.149–74.

Veilleux, N., Shattuck-Hufnagel, S. and Brugos, A. (2004), 'The new English ToBI homepage'. **URL:** *http://anita.simmons.edu/~tobi/index.html (Feb 2011)*

Vizcaino-Ortega, F. (2002), "A preliminary analysis of yes/no questions in Glasgow English", *in Proceedings of the Speech Prosody 2002 Conference*, Aix-en-Provence, pp. 683–86.

Wagner, R. A. and Fischer, M. J. (1974), "The string-to-string correction problem", *Journal of the Association for Computing Machinery*, 21(1), pp.168–73.

Warnow, T. (2009), Computational methods for inferring evolutionary histories of languages. Keynote address presented at The Swadesh Centenary Conference, Max Planck Institute for Evolutionary Anthropology, Leipzig.

Warren, P. (2005), "Patterns of late rising in New Zealand English: Intonational variation or intonational change?", *Language Variation and Change*, 17, pp.209–30.

Warren, P. and Daly, N. (2005), "Characterising New Zealand English intonation", *in* A. Bell, R. Harlow and D. Starks, eds, *Languages of New Zealand*, Victoria University Press, Wellington, pp. 217–37.

Welby, P. (2003), The slaying of Lady Mondegreen, being a study of French tonal association and alignment and their role in speech segmentation, PhD thesis, The Ohio State University, Columbus, OH.

Welby, P. and Loevenbruck, H. (2006), "Anchored down in Anchorage: Syllable structure and segmental anchoring in French", *Italian Journal of Linguistics*, 18, pp.74–124.

Wells, B. and Peppé, S. (1996), "Ending up in Ulster: Prosody and turn-taking in English dialects", *in* E. Couper-Kuhlen and M. Selting, eds, *Prosody in conversation*, CUP, Cambridge, pp. 101–30.

Wells, J., ed. (1982), *Accents of English*, CUP, Cambridge.

Xu, Y. (1998), "Consistency of tone-syllable alignment across different syllable structures and speaking rates", *Phonetica*, 55, pp.179–203.

Xu, Y. (2001), "Fundamental frequency peak delay in Mandarin", *Phonetica*, 58, pp.26–52.

Xu, Y. (2005), "Speech melody as articulatory implemented communicative functions", *Speech Communication*, 46(3-4), pp.220–51.

Xu, Y. and Sun, X. (2002), "Maximum speed of pitch change and how it may relate to speech", *Journal of the Acoustical Society of America*, 111(3), pp.1399–1413.

Yip, M. (2002), *Tone*, CUP, Cambridge.

Yuan, J. (2006), "Mechanisms of question intonation in Mandarin", *in ISCSLP LNAI 4274*, pp. 19–30.

Yuan, J., Shih, C. and Kochanski, G. P. (2002), "Comparison of declarative and interrogative intonation in Chinese", *in Proceedings of Speech Prosody 2002*, Aix-en-Provence, pp. 711–4.