THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Using Corpora of Recorded Speech for Historical Phonology

Citation for published version:
Maguire, W 2014, Using Corpora of Recorded Speech for Historical Phonology. in P Honeybone & J Salmons (eds), The Oxford Handbook of Historical Phonology. Oxford University Press, pp. 164-172. DOI: 10.1093/oxfordhb/9780199232819.013.035

Digital Object Identifier (DOI):
10.1093/oxfordhb/9780199232819.013.035

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
The Oxford Handbook of Historical Phonology

OPEN ACCESS

Oxford Handbooks Online

Using Corpora of Recorded Speech for Historical Phonology 🔓

Warren Maguire

The Oxford Handbook of Historical Phonology
Edited by Patrick Honeybone and Joseph Salmons

⊟ Abstract and Keywords

This chapter examines the use of electronic corpora of recorded speech in historical phonological analysis. On the one hand, corpora of this kind provide us with unique insights into the synchronic phonologies of varieties, allowing us to analyse change in real and apparent time in ways which have never been possible before. On the other hand, the fact that these corpora typically provide us with fragmentary, production-only data which are rarely strictly comparable with each other means that there are considerable limitations on our phonological interpretation of them.

Keywords: electronic corpora, synchronic phonology, production, strict comparability, fragmentary data

1 Introduction

How were the Proto-Indo-European stops (never mind the laryngeals) really pronounced? What was the nature of Old English ‹æ›, ‹ea›, and ‹ēa›, and how were they distinguished? What was 'so scharp, slytting and frotyng and unschape' about northern Middle English dialects? Was there a difference between the vowels in the MEAT and MATE lexical sets in sixteenth-century London English and, if so, what was it? How did Ben Jonson pronounce /r/, 'the dog's letter', and what was the nature of the allophonic variation he seems to describe? These questions (for background to some of the questions, see the contributions by Kümmel, Minkova, and Gordon in this volume) and many others like them from the history of English and other languages could easily be answered if we had audio recordings of speakers from those periods of history.

But of course audio recordings only began to be produced in the second half of the nineteenth century (Morton 2004: 2–7), and recordings of speech didn't become common until the early twentieth century (the British Library's online archive of *Early Spoken Word Recordings* in English contains samples going back as far as 1905). Even so, they have transformed the way we think about language and how we analyse phonology, synchronic and diachronic. This chapter discusses not only the (obvious) advantages of using corpora of recorded speech for historical phonological analysis but also some of the problems we encounter when we attempt to do so. Its focus is on English, but the same issues apply to any language. For the purposes of this chapter, a corpus of audio recordings of speech may be defined as 'a body of language data which can serve as a basis for linguistic analysis and description' (Bauer 2002: 98) made up of audio recordings, analogue or digital (see Bauer 2002 and D'Arcy 2011 for further discussion of the meaning of the term 'corpus' and for excellent overviews of different types, examples, and uses of linguistic corpora).

2 Advantages

Although audio recordings are only available for the last few generations of speakers, it is still possible to investigate language change in real time (see Bowie and Yaeger-Dror, this volume) using them. Thomas (2001) is an excellent illustration of what we can learn by comparing audio recordings from the early twentieth century (of North American speakers born in the second half of the nineteenth century) with recordings of speakers from the same localities at the end of the twentieth century. The changes he identifies in vowel systems are often striking, and Thomas's study shows that the history of recorded speech is already long enough for us to study language change in real time. Other instructive examples of such research include Purnell et al. (2005a,2005b), who analyse fortition of word-final stops in Wisconsin English in real time, and Wagener (1997), who analyses changes in German in the United States using recordings from two different time periods. Another striking example of this is the examination of change in the pronunciation of /r/ in Montreal French, using trend and panel studies, reported in Sankoff and Blondeau (2007). It is only through the use of recorded corpora that we can see change, or lack of change, during the lifespan in such detail, and investigate such crucial questions in historical phonology at all.

Given the short history of recording technology, and the difficulty of finding comparable recordings from different time periods (see Section 4), real-time studies are often not possible. But corpora of recorded speech are ideal for studying language change in 'apparent time' (see Bailey 2002). By analysing the speech of comparable speakers from different age cohorts, we can, with some reservations, gain insights into language change even when we have no other evidence. Thus, for example, a number of apparent time studies of a corpus of Tyneside English from the 1990s (the *Phonological Variation and Change in Contemporary British English* [PVC] corpus) have revealed that the variety has been subject to a range of exogenous changes such as glottal replacement, change in the realization of /r/, and levelling of localized vowel pronunciations (see, for example, Milroy et al. 1994, Watt and Milroy 1999, and Foulkes and Docherty 2000). Highly localized patterns are most characteristic of the speech of older working-class males, whilst patterns which are known to be spreading through varieties of British English are most characteristic of young female speech.

Even with only an hour-long recording of someone's speech, we can, to a reasonable extent, identify the likely phonemic distinctions they have, the typical realizations of these phonemes, their major allophones (including cross word-boundary phenomena), and get a good indication of the lexical distribution of these phonemes. When we analyse samples of the speech of a group of speakers who have roughly the same linguistic system, we can add even more detail to the picture, since what we may be lacking for one speaker might be present for another, giving us a wider indication of the parameters of variation in the speech community (even if we cannot assume that all of this variation is present in the speech of every individual).
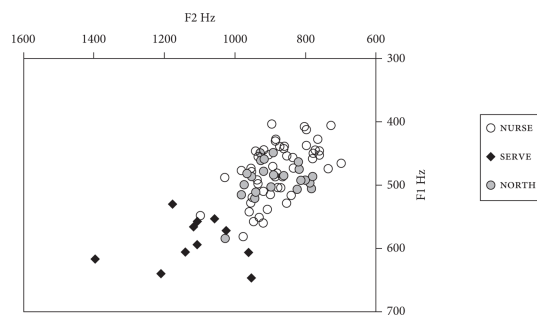
Click to view larger

Figure 1 . The NURSE, SERVE, and NORTH lexical sets for a working-class male, born 1930s
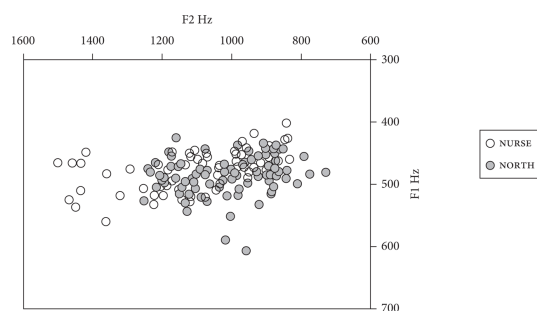
Click to view larger

Figure 2 . The NURSE and NORTH lexical sets for a working-class male from Tyneside, born 1920s

This information can be correlated with geographic, social, and historical factors which help us to explain why these individuals' phonologies are the way they are. Two areas in historical phonology which have benefitted enormously from the study of audio recordings are chain shifts and mergers (see Gordon, this volume, for a detailed overview). Relying as we must on written evidence, it is almost impossible to see the English Great Vowel Shift in action—no matter how good the evidence from spelling, rhymes, puns, and metalinguistic comment, it can never give us an up-close view of how a chain shift actually happens (but see the chapters by Lass, Minkova, and Unger in this volume for how far we can push orthographic evidence in this kind of analysis). With audio recordings, we can see in minute detail how each phoneme is pronounced, the degree of variation in their pronunciation, how close they are in phonetic space, what phonetic factors are involved in maintaining distinctions between them, and the social distribution of the change. Labov (1994: 113–291) is a groundbreaking study of what we can learn about chain shifts from examining corpora of recorded speech. So although we will never be able to investigate the Great Vowel Shift in the same way, we can learn from changes such as the Northern Cities Shift about the forces and principles that were likely to have been in operation in earlier stages of the history of the phonology of English and other languages. Likewise, the availability of audio recordings has revolutionized the study of phonemic mergers (see Labov 1994: 295–418 for detailed discussion). Take the example of the 'NURSE–NORTH Merger' in Tyneside and Northumberland English (see Wells 1982: 374–5, Watt 1998a, 1998b, and particularly Maguire 2008). This reported merger, under an [ɔː]-type vowel, is amply attested (for what that is worth) in written representations of the dialect and in phonetic transcriptions in traditional dialect studies. But nowadays it is very much restricted to the speech of older working-class males, and is only variable even there, suggesting that it is disappearing from the speech community. How did this apparent reversal of the merger come about, and might it be the case that this was never a merger in the first place? Fortunately we have several corpora of recorded speech from the northeast of England in the twentieth century which allow us to explore this phenomenon in detail. In particular, the recordings produced by the *Tyneside Linguistic Survey* (TLS; Pellow et al. 1972; Allen et al. 2006[1]) in the early 1970s are a rich source of information on vernacular Tyneside speech. A detailed phonetic analysis of these recordings (Maguire 2008) reveals a complex picture of variation, with some speakers having no merger of the NURSE and NORTH lexical sets, others having complete merger (in production at least—see Section 3), and still others having some degree of overlap in the phonetic realizations of the two vowels (from slight overlap to almost, but not quite, complete phonetic identity). The picture which emerges is much more complex than that available from other sources—the existence of the merger is confirmed, but only for some speakers, and the existence of substantial overlap but statistical non-identity of the two vowels suggests that the merger may never have been complete for many speakers, or may be being reversed in a rather subtle fashion in the speech community. Figures 1 and 2 reveal the kinds of complex patterns that are still retrievable from these legacy recordings, which were made at a time when acoustic analysis was difficult and not commonly done, and which have degraded quite considerably since they were made. The speaker in Figure 1 (working-class male, born 1930s) has a merger in production of the NURSE and NORTH vowels on both the F1 and F2 dimensions (i.e. there is no statistical difference between the distributions of the two vowels), but retains a significant distinction between these and the SERVE subset of NURSE (i.e. those words which had late Middle English /er/). This is exactly the pattern we would expect in the area from traditional dialect studies such as the *Survey of English Dialects* (SED; Orton and Dieth 1962–71), and it is reassuring to see an acoustic analysis of this TLS recording confirm the pattern identified by the impressionistic traditional dialect analyses. The speaker in Figure 2 (a working-class male, born 1920s), on the other hand, does not distinguish SERVE from the rest of NURSE but does have statistically distinct NURSE and NORTH lexical sets (on both F1 and F2 dimensions). However, the two lexical sets overlap to a large extent and are in a situation of near merger.

It is noteworthy that it is really only through phonetic analysis of recorded speech that the important phenomenon of near merger (Gordon, this volume) can be identified, and the existence of the substantial TLS corpus allows us to identify another example of this in Tyneside English.

3 Limitations

It is clear, then, that corpora of recorded speech are an enormous boon for anyone working on historical phonology, and where they are available historical phonologists can hardly fail to take them into account. Despite the obvious benefits of corpora of recorded speech, however, I wish to highlight some rather important problems

here, too.

Most importantly, and something which the richness of data available to us in corpora may almost make us forget, is the fact that speech is not phonology. Whether we envisage phonology as an abstract system rather indirectly related to speech (e.g. Hale and Reiss 2008 and Hale et al., this volume) or as a malleable cognitive state which is in a constant feedback loop with what we say and hear (e.g. Bybee 2001; Bybee, this volume), speech is an articulatory and acoustic phenomenon whilst phonology is mental representation. Even if we lose the ability to hear or speak (e.g. through deafness, expressive aphasia, or dysarthria), we do not lose our knowledge of our native language's sound patterns. So what we capture in corpora is not phonology either—but rather electronic or magnetic encodings of speech sound waves. Although speech is our main body of *evidence* for phonology, we must always remember that there are radical differences between the two (see Ritt 2004: 3–7), and some of these come sharply into focus when we attempt phonological analysis on the basis of audio recordings.

One problem is that a recording of a person's speech can only ever be a snapshot of their linguistic behaviour, no matter how long the recording. Assuming that you live for seventy years (and many people in the Western world are living much longer), you'll have been alive for about 613,600 hours. Even if the time you spend speaking (and listening) is only a fraction of this, and even if our speech is endlessly repetitive (at the phonetic and phonological levels at least), it is obvious that an hour-long recording of your speech represents only a tiny fragment of your linguistic repertoire.

Even so, corpora of (relatively) short recordings have proven to be extremely useful for analysing a wide range of phonetic and phonological features, as the vast body of sociophonetic work illustrates, for example. But recorded corpora are less useful when we want to understand the nature of complex phonological phenomena. Good examples of this are T-to-R (Buchstaller et al. 2013), found especially in northern English Englishes, and the Scottish Vowel Length Rule (Aitken 1981). T-to-R involves alternation between /t/ and /r/ in word-final, typically foot-internal position, as in *hit* [hɪt], *hit it* [ˈhɪɹɪt], and *lot* [lɒt], *lot of* [ˈlɒɹə]. This phenomenon is readily apparent in recordings of speakers from northern England (see Docherty et al. 1997), but the precise conditions for the alternation are very difficult to define, involving as they do the nature of the preceding vowel, metrical position, word frequency, and lexicalization. Although recordings will reveal that some words in some environments will be subject to T-to-R, they will not tell us all of the words that can be affected, how often they are affected, all of the metrical constraints on the phenomenon, and, least of all, which words cannot be affected (since absence of evidence is not evidence of absence, certainly not in a relatively small sample). Even hundreds of hours of recordings of a range of speakers might not reveal the full picture, and it is no wonder that other techniques have been explored for investigating T-to-R in these dialects (Buchstaller et al. 2013). Similarly, the Scottish Vowel Length Rule (SVLR) involves alternations in vowel length and quality dependent upon following phonemes, morpheme boundaries, and, in all likelihood, metrical position and lexical or frequency affects. Capturing a substantial part of this in an audio recording is unlikely (not to mention the difficulty of determining vowel length rules in short samples of speech where other factors, such as speech rate, obscure the picture even further). Just to give a concrete example, Table 1 summarizes the numbers of words and tokens in the PRICE lexical set for two older working-class male speakers (J34 and F35) of Tyneside English from the PVC corpus. This rich corpus of vernacular Tyneside speech from the mid 1990s reveals that the PRICE vowel is subject to SVLR conditioning in Tyneside English (Milroy 1995), such that [aɪ]-type allophones are found before voiced fricatives, schwa (including [ə] derived from earlier /r/ in this non-rhotic variety) and morpheme boundaries, and an [ɛi]-type diphthong is found elsewhere. In approximately 45 minutes of conversation with each other, J34 and F35 produced over 4000 words each, and 169 (56 words) and 326 (46 words) PRICE tokens respectively. This looks like a good number of tokens for analysis, although there are very few tokens for some environments (e.g. before voiced fricatives and before schwa). But when we divide the tokens up into individual speakers, lexical items, historical lexical sets, and phonological environments, things look less promising. Thus, the word *aye* ('yes'), which may well be phonologically atypical, is overwhelmingly common. If we remove it from the analysis (resulting in the bracketed figures in Table 1), the number of tokens is substantially reduced (116 and 134), and data for the pre-morpheme boundary environment is suddenly very sparse. This is especially the case for words where the vowel is morpheme-final but followed by past tense/past participle *-(e)d* (i.e. four items in total, *died, terrified, tied, tried*) which are crucial for understanding the role of morphology in SVLR conditioning (since they reveal whether or not it is morpheme boundaries rather than word boundaries which condition the vowel alternation). Additionally, some of the words where the vowel is followed by a morpheme boundary belong to a different historical lexical set (e.g.

die, fly, high) which may pattern differently (with final [εi] rather than [aɪ]) in modern Tyneside English (see Milroy 1995), further complicating the picture. All of this means that it may be impossible to draw any significant conclusions about the pronunciation of the PRICE vowel in morpheme-final position in the speech of these two individuals, who constitute exactly half of the older working-class male cohort in the corpus, even assuming that the tokens involved are of sufficient quality to allow meaningful phonetic analysis.

Table 1. Numbers of PRICE words and tokens for two speakers in the PVC corpus						
	J34		**F35**		**Total**	
Environment	**Words**	**Tokens**	**Words**	**Tokens**	**Words**	**Tokens**
+ voiceless C	18	39	18	90	27	128
+ voiced fricative	4	8	5	9	7	17
+ other voiced C	14	38	16	26	23	65
+ schwa	5	9	2	3	7	12
+ morpheme boundary	15 (14)	75 (22)	5 (4)	198 (6)	16 (15)	273 (28)
Total	56 (55)	169 (116)	46 (45)	326 (134)	80 (79)	495 (234)

These data raise another issue which is problematic for anyone using recorded audio corpora for historical phonological analysis. In addition to words such as *die, fly*, and *high* belonging to a different historical lexical set, other PRICE words belong to different historical sets again. Words spelt *-ight* typically derive from Middle English /ixt/, and are still sometimes pronounced with an [iː]-type vowel in Tyneside English, whilst words spelt with *-ind* typically derive from Middle English /ind/, and may still occasionally be heard with [ɪ]. These pronunciations are so deeply embedded in the vernacular that they rarely occur in recorded speech, but are often known even when they are not used. This means that there is a significant disjunction between production and knowledge that cannot be bridged by relying on recorded corpora alone.

For instance, you might talk to (and record) me for hours without realizing that I have, as part of my phonology, a variable distinction between MEET [i] and MEAT [i]~[e] (which may in turn, depending upon phonological environment and probabilistic factors, be different than MATE [ɪə]; see Milroy and Harris 1980). Unless I'm talking to someone with a similar accent and social background to myself in a very informal situation, I'm unlikely to use the lower variant of MEAT at all (which I only use some of the time even then), and it would take many hours of recordings of such speech to gain an appreciation of the lexical, phonological, and phonetic characteristics of it. Conversely, acoustic analysis of the recordings from the TLS (Maguire 2008) suggests that some speakers have a very localized merger of the NURSE and NORTH lexical sets. But we have no idea what was going on inside these speakers' heads—did they know there was a possible (more standard) distinction, and could they have produced it if put on the spot (e.g. in a reading task)? One suspects that some might have been able to do so, which puts a question mark over the whole idea that these speakers have a merger at all (since a merger in production but not in perception can hardly be described as a phonological merger). So what we hear is, in all likelihood, a representation of only a fragment of what the speakers know, and this is crucial for understanding things such as splits, mergers, and lexical distributions of phonemes.

4 Methodological Considerations

These problems relate to a number of methodological issues which affect our understanding of the speech that we analyse, and it is clear that using recorded corpora for historical phonology is far from straightforward (even assuming the recordings we have are of sufficient quality to make analysis worthwhile, which may not be the case

for old legacy corpora). The 'Observer's Paradox' (Labov 1966: 86) is a well-known effect which can influence the kind of speech we record in the first place, so that the material we are working with may not be wholly representative of the vernacular we hope to study. Furthermore, the speakers and styles of speech we have recorded may not be strictly comparable with other corpora of recordings or with other kinds of data. Thus, the PVC (mid 1990s) consists of a stratified sample of speakers from the West End of Newcastle upon Tyne in naturalistic dyadic conversations, the TLS (early 1970s) is a collection of one-to-one semi-formal interviews with a fairly random sample of residents from Gateshead (including some non-natives), the SED audio recordings (mid 1950s) from the area are one-to-one interviews with older male manual labourers from villages surrounding Tyneside who were chosen on the basis that their speech was particularly old-fashioned and divergent from Standard English, and the SED phonetic transcriptions represent the interpretations of the fieldworker attempting to record only the most localized forms of speech used by these same individuals. Beyond that, we have a substantial body of nineteenth-century dialect literature from the Tyneside area which represents the attempts of a diverse range of writers to capture the essence of the dialect orthographically. We shouldn't expect these bodies of data to have recorded the same kinds of speech, with the result that differences between them may be attributable to a range of factors other than language change.

It is not surprising, then, that researchers analysing phonology using corpora of recorded speech often use additional techniques to supplement the information they get from them. These approaches include targeted elicitation tasks such as word-lists and reading passages (e.g. Trudgill 1974), probing of speakers' knowledge of phonological patterns through direct questioning (e.g. Maguire et al. 2010; Buchstaller et al. 2013), judgement tasks such as minimal pair and commutation tests (see Labov 1994: 353–7), and perception experiments such as the Coach Test (Labov 1994: 403–6).

But the importance of recorded corpora of speech cannot be overstated. Without them our insights into synchronic and diachronic phonology are severely limited. We can't discover everything about a person's phonology or about the shared properties of the phonologies of groups of speakers, but corpora are a very good place to start and give us much that we might want to know. Furthermore, they introduce a strong element of objectivity and accountability to the data which is not possible with intuitions and general observations. Their importance for historical phonology is only going to increase as the history of recorded corpora lengthens.

Bibliography

Aitken, A. J. 1981. The Scottish vowel-length rule. In Michael Benskins and Michael Samuels (eds), *So Many People, Longages and Tonges: Philological Essays in Scots and Medieval English Presented to Angus McIntosh*, 131–57. Edinburgh: Edinburgh University Press.

Allen, Will, Joan Beal, Karen Corrigan, Warren Maguire, and Hermann Moisl. 2006. A linguistic 'time capsule': The Newcastle Electronic Corpus of Tyneside English. In Joan Beal, Karen Corrigan, and Hermann Moisl (eds), *Creating and Digitizing Language Corpora, vol. 2: Diachronic Databases*, 16–48. Houndmills, Basingstoke: Palgrave Macmillan.

Bailey, Guy. 2002. Real and apparent time. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change*, 312–22. Oxford and Malden: Blackwell.

Bauer, Laurie. 2002. Inferring variation and change from public corpora. In J. K. Chambers, Peter Trudgill, and Natalie Schilling-Estes (eds), *The Handbook of Language Variation and Change*, 97–114. Oxford and Malden: Blackwell.

British Library. *Early Spoken Word Recordings* archive. Available online at <>, accessed 26 October 2012.

Buchstaller, Isabelle, Karen Corrigan, Anders Holmberg, Patrick Honeybone, and Warren Maguire. 2013. T-to-R and the Northern Subject Rule: Questionnaire-based structural, geo-, and sociolinguistics. *English Language and Linguistics,* 17: 85–128.

Bybee, Joan. 2001. *Phonology and Language Use*. Cambridge: Cambridge University Press.

D'Arcy, Alexandra. 2011. Corpora: Capturing language in use. In Warren Maguire and April McMahon (eds),

Analysing Variation in English, 49–72. Cambridge: Cambridge University Press.

Docherty, Gerard J., Paul Foulkes, James Milroy, Leslie Milroy, and David Walshaw. 1997. Descriptive adequacy in phonology: A variationist perspective. *Journal of Linguistics,* 33: 275–310.

Foulkes, Paul and Gerard J. Docherty. 2000. Another chapter in the story of /r/: 'labiodental' variants in British English. *Journal of Sociolinguistics,* 4: 30–59.

Hale, Mark and Charles Reiss. 2008. *The Phonological Enterprise*. Oxford: Oxford University Press.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, William. 1994. *Principles of Linguistic Change, vol. 1: Internal Factors.* Oxford and Cambridge, MA: Blackwell.

Maguire, Warren. 2008. What is a merger, and can it be reversed? The origin, status and reversal of the Nurse–North Merger in Tyneside English. PhD thesis, Newcastle University.

Maguire, Warren, April McMahon, Paul Heggarty, and Dan Dediu. 2010. The past, present and future of English dialects: Quantifying convergence, divergence and dynamic equilibrium. *Language Variation and Change,* 22(1): 1–36.

Milroy, James. 1995. Investigating the Scottish vowel length rule in a Northumbrian dialect. *Newcastle and Durham Working Papers in Linguistics,* 4: 187–96.

Milroy, James and John Harris. 1980. When is a merger not a merger? The meat/mate problem in a present-day English vernacular. *English World-Wide,* 1: 199–210.

Milroy, James, Lesley Milroy, Sue Hartley, and David Walshaw. 1994. Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language Variation and Change,* 6: 327–57.

Morton, David. 2004. *Sound Recording: The Life Story of a Technology*. Westport, CT: Greenwood Press.

Orton, Harold and Eugen Dieth. 1962–71. *Survey of English Dialects: Basic Material*. Leeds: University of Leeds Press.

Pellowe, John, Barbara Strang, Graham Nixon, and Vincent McNeany. 1972. A dynamic modelling of linguistic variation: The urban (Tyneside) linguistic survey. *Lingua,* 30: 1–30.

Purnell, Thomas, Joseph Salmons, Dilara Tepeli, and Jennifer Mercer. 2005a. Structured heterogeneity and change in laryngeal phonetics: Upper Midwestern final obstruents. *Journal of English Linguistics,* 33(4): 307–38.

Purnell, Thomas, Joseph Salmons, and Dilara Tepeli. 2005b. German substrate effects in Wisconsin English: Evidence for final fortition. *American Speech,* 80(2): 135–64.

Ritt, Nikolaus. 2004. *Selfish Sounds and Linguistic Evolution*. Cambridge: Cambridge University Press.

Sankoff, Gillian and Hélène Blondeau. 2007. Longitudinal change across the lifespan: /r/ in Montreal French. *Language,* 83(3): 560–88.

Thomas, Erik. 2001. *An Acoustic Analysis of Vowel Variation in New World English*. Durham, NC: Duke University Press.

Trudgill, Peter. 1974. *The Social Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

Wagener, Peter. 1997. Nach 40 Jahren: Zu individuellen Veränderungen der gesprochenen Sprache. In Gerhard Stickel (ed.), *Varietäten des Deutschen: Regional- und Umgangssprachen*, 291–307. Berlin: Mouton de Gruyter.

Watt, Dominic. 1998a. Variation and change in the vowel system of Tyneside English. PhD thesis, Newcastle University.

Watt, Dominic. 1998b. One vowel short: The (reported) merger of NURSE and NORTH in Tyneside English. In Claude Paradis, Diane Vincent, Denis Deshaies, and Marty Laforest (eds), *Papers in Sociolinguistics: NWAVE-26 à l'Université Laval*, 149–58. Québec: Éditions Nota bene.

Watt, Dominic and Lesley Milroy. 1999. Patterns of variation and change in three Newcastle vowels: Is this dialect levelling? In Paul Foulkes and Gerard J. Docherty (eds), *Urban Voices: Accent Studies in the British Isles*, 25–47. London: Arnold.

Wells, J. C. 1982. *Accents of English, 3 vols.* Cambridge: Cambridge University Press.

Notes:

([1]) The PVC and TLS recordings are available online: http://research.ncl.ac.uk/necte/.

Warren Maguire

Warren Maguire works in the department of Linguistics and English Language at the University of Edinburgh. His research is focussed on variation and change in the phonology of regional dialects of English and Scots in Britain and Ireland. He has recently published articles on Pre-R Dentalisation in northern England and on Alexander J. Ellis's 'The Existing Phonology of English Dialects'. He is co-editor of Analysing Variation in English (2011, CUP), and of a special issue of English Language and Linguistics on phonological mergers (2013).