*English Language and Linguistics* 17.2: 325–357. © Cambridge University Press 2013
doi:10.1017/S1360674313000051

# New contrast acquisition: methodological issues and theoretical implications

JENNIFER NYCZ
*Georgetown University*
(Received 8 May 2012; revised 15 February 2013)

This article presents data on the acquisition of the low back vowel contrast by native speakers of Canadian English who have moved as adults to the New York City region, examining how these speakers who natively possess a single low back vowel category have acquired the low back vowel distinction of the new ambient dialect. The speakers show remarkable first dialect stability with respect to their low back vowel system, even after many years of new dialect exposure: in minimal pair contexts, nearly all of the speakers continue to produce and perceive a single vowel category. However, in word list and conversational contexts, the majority of speakers exhibit a small but significant phonetic difference between words like *cot* and *caught*, reflecting the separation of these word classes in the new dialect to which they are exposed; moreover, the realization of these words shows frequency effects consistent with a lexically gradual divergence of the two vowels. These findings are discussed in terms of their implications for theories of phonological representation and change, as well as their methodological implications for the study of mergers- and splits-in-progress.

## 1 Introduction

The opposite of merger is phonemic split: when one category becomes two, either in a language variety or in the phonological system of an individual. Splits are not as well studied as mergers, probably because they are less often observed; phonological mergers tend to spread at the expense of distinctions (Herzog 1965; Labov 1994), a dialectological finding so robust that it has been given a name, 'Herzog's Principle', after Herzog's study of mergers affecting high vowels in the Yiddish of northern Poland. Yet both types of change touch on central theoretical and methodological questions in phonology, language change and the intersection of these two areas: what kind(s) of knowledge do speakers have about the sounds of their language? In what ways does this knowledge reflect variation and change in the community? How do we investigate, characterize and formalize individual speaker knowledge in light of community variation?

In this article I will review some of the specific methodological and theoretical issues that have been raised by the study of mergers, then describe how the study of splits can shed further light on these concerns. I will then present the results of a sociolinguistic study of mobile adults – Canadians in the New York region – who show evidence of acquiring a low back vowel split as a result of dialect contact, and discuss the theoretical and methodological implications of these findings.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

### *1.1 Methodological issues in the study of contrast and merger*

Of all the various types of sound change, mergers and splits are particularly interesting from a phonological perspective because they involve a change in the number of contrastive elements within a language. Speech sounds (or more abstractly, *phonemes*) do not bear referential meaning, but serve as the building blocks from which meaningful units (*morphemes*) can be composed, and by which meaningful units can be distinguished from one another. A phoneme's principal job, in other words, is to contrast with other phonemes. A core part of phonological knowledge is knowing what these contrastive elements are.

How do we identify the contrastive elements of a speaker's language given all the phonetic variation which characterizes its surface forms? The classic method for uncovering contrast is the minimal pair test. In a fieldwork context, the linguist can present a speaker with two strings that differ in just one sound (e.g. [pat] and [p$^{h}$at]). The speaker is then asked to say whether these strings are instances of the same word or (potentially) different words.[1] This minimal pair judgment reveals whether a given difference in sound can be used to make a difference in meaning for the speaker, and thus whether it is contrastive.[2]

Such clear-cut results are probably the norm in cases where the community variety is not undergoing any changes with respect to the sounds of interest. The situation can become more complicated, however, when the community variety is characterized by a merger-in-progress that destabilizes the relationship between these sounds. The existence of *near-mergers* (Labov *et al.* 1991) in such contexts has been revealed through the use of minimal pair tests, though it is important to note that these are used by sociolinguists in a very different way from how they might be used by fieldworkers attempting to discover the phonemic inventory of a language. Rather than starting out with two strings of sounds that differ in one segment, and then asking whether these can be two different words, the sociophonetician will present the speaker with two different words printed in standard language orthography (e.g. 'cot' and 'caught'), then ask the speaker to say these words out loud and judge whether they sound the same. This task thus elicits information about two types of speaker knowledge: implicit knowledge regarding *how* these forms are produced and explicit knowledge *that* two forms are different (or the same).

[1] At least, this is how things are purported to work, though this knowledge seems limited to the linguistics oral tradition. Labov (1994) comments that he 'has not found any detailed descriptions of minimal pair tests from the period of structural linguistics, when methods for describing languages were prominently discussed' (353). Ladefoged's (2003) guide to fieldwork notes the usefulness of minimal pairs in uncovering the contrastive elements of a language, not as a task which elicits speaker intuitions about contrast, but as a later analytic tool which can be used on already collected data (the real world version of a phonology class phonemicization problem set). Vaux & Cooper's (1999) fieldwork guide does not mention minimal pairs at all in the chapter on segmental phonology, perhaps due to their view that informant intuitions about the sound patterns of their language are unreliable (79).

[2] A positive result from a minimal pair test simultaneously demonstrates another feature of contrastiveness: if two sounds contrast, the presence of one rather than the other cannot be predicted by phonological environment.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

In many cases, these two types of knowledge will align in expected ways: speakers will produce a clear difference and also acknowledge it, or produce the relevant pairs as homophones and accordingly judge them to sound the same. However, mismatches between production and perception[3] can also occur. Sometimes a speaker will produce no difference, but claim that one exists; this probably can be explained in terms of the influence of orthography and a belief that things that are spelled differently sound different (Labov *et al.* 1991). In other cases, a speaker will consistently produce a small measurable phonetic difference between relevant words across pairs, but claim that the pairs sound the same. A well-known example is Bill Peters, an older speaker from central Pennsylvania whom Labov interviewed in 1970. Bill Peters produced a small but consistent difference between words like *cot* and *caught* in minimal pair tests, but 'never hesitated' (Labov 1994: 364, fn. 10) in judging such pairs to be homophonous.

Studies of speakers whose community varieties are characterized by mergers-in-progress thus highlight the importance of asking the right questions when attempting to access speakers' knowledge of the sounds of their language. In the first kind of minimal pair test described above, only intuitions – knowledge-*that* – are probed. In the second kind of minimal pair test, both productions and intuitions about these productions are queried, in some cases revealing a dissociation between the knowledge *that* sounds are a certain way and the knowledge of *how* those sounds are produced.

Of course, sociolinguistic studies of merger-in-progress draw on more types of data than minimal pair tests. Speakers may vary in how they behave across different types of language tasks, indicating that the simple distinction between knowledge-*that* and knowledge-*how* made so far is not sufficient to capture the complexity of facts that speakers internalize regarding the sounds in their language. For example, though Bill Peters showed a near-merger of *cot/caught* in minimal pair context, he produced a clear distinction between relevant words in his spontaneous speech. In a similar well-documented individual case, Dan Jones of Albuquerque (Labov *et al.* 1972) produced *no* distinction between words like *pool/pull*, *fool/full* in minimal pair tests, and accordingly judged such pairs to sound the same. In tokens of the same words produced for a commutation task and in interview context, however, Dan seemed to produce a clear distinction.

Such differences across task types show that a speaker's knowledge-*that* two sounds are the same or different is not straightforwardly derived from the magnitude of the phonetic difference between these sounds in the speech of that speaker generally, but must reflect some other norm. At the same time, when knowledge-*that* is explicitly queried in a minimal pair test, it mediates the usual course of knowledge-*how*, phonetically neutralizing (or nearly so) a contrast which is otherwise clearly made. Labov *et al.* note that in controlled styles such as minimal pair readings 'many of the important allophonic differences are wiped out, and, depending on the particular

[3] A reviewer points out that the 'perception' portion of the minimal pair test is more accurately described as an introspection task. I retain the language of 'perception results' and 'merger-in-perception' here, because these are the terms used in the literature on near-merger (e.g. Labov *et al.* 1991).

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

sociolinguistic configuration, the mean values may shift radically backwards towards an older, corrected value, or radically forwards towards the apparent target of the change'. (Labov *et al.* 1991: 57). An example of the first option is described in Johnson's (2010) study of low back merger among children whose parents have a low back vowel distinction, but who live in areas of New England which are increasingly characterized by merger. Such children tend to be merged in their spontaneous speech, reflecting the patterns of their relatively recently acquired peer group, but produce a distinction in more controlled styles, reflecting the older norm learned from their non-merged parents. Bill Peters and Dan Jones exemplify the second option: both men reflect the incoming community norm in their minimal pair judgments and productions, even though the change has not generally come to characterize their own speech.

Studies of mergers-in-progress in sociophonetics have thus shown that different tasks can reveal complex relationships between different types of knowledge that speakers have about the sounds of their language. They may implicitly know that two word classes are produced differently (as indicated by their spontaneous speech), but at the same time seem to know that these word classes ought to sound the same, reflecting wider community norms.

### *1.2 Theoretical relevance of mergers and splits*

A discussion of how to uncover a speaker's knowledge of the sounds of their language naturally raises the question of what form this knowledge takes. While there are many specific theories regarding the nature of phonological representations and how they map onto surface forms, these diverse views can generally be divided into two major groups: *abstractionist models* and *phonetically rich models*. I begin this section with an overview of each of these approaches, moving on to a discussion of how phonological contrast and near-merger is modeled in each type of framework. Finally, I will explain how the study of splits can help to decide between these two views.

#### *1.2.1 Abstractionist view of representation*

The mainstream view in phonology is that underlying representations are quite abstract compared to surface forms. This view of representation has a long history in phonological thought, being a principal component of structuralism (Saussure 1916) and the linguistic theories of the Prague School (Trubetzkoy 1969 [1939]; Jakobson 1962), and was further articulated in *The Sound Pattern of English* (Chomsky & Halle 1968). Later developments within generative theory such as autosegmental phonology (Goldsmith 1979) and feature geometry (Clements 1985; Clements & Hume 1995) made the underlying representations more complex, but continued to hew to the same principle of abstractness from surface form. More currently, analyses carried out within Optimality Theory typically assume abstract, feature-based representations (e.g. Kager 1999).[4]

[4] Optimality Theory itself is agnostic regarding the form of underlying representations. OT analyses can in theory be carried out on a variety of representational types – see e.g. Gafos (2002), who builds an OT grammar that operates on gestural coordination schemes.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

In the abstractionist view, representations are minimally specified, containing only the information needed to differentiate all phonemes in the inventory.[5] This information takes the form of features which are given phonetically inspired names such as [voice] or [nasal]; it is important to remember, however, that these labels are essentially mnemonic, as the real purpose of features is to distinguish phonemes from one another. No phonetic information is present in the underlying representation; the articulatory and/or acoustic spelling out of these labels is determined by phonetic implementation rules after the derivation of surface forms.

As with most ideas in linguistics, this notion of abstractness comes bundled in a larger theoretical package of interconnecting assumptions and principles. Closely intertwined with the idea of minimally specified underlying representations is the assumption that there is typically only one such representation per lexical item,[6] from which all surface variation derives. Because these representations do not reflect surface variation, they are stable over time, though they can in principle change via the addition, subtraction, or alteration of one or more features. These unique, minimally specified underlying representations serve as the input to phonological rules which alter features of the representation to produce intermediate and ultimately surface forms. Representations and rules are distinct components in this view. Underlying representations contain all – and only – that which is arbitrary and unpredictable about the word form (such as segment order and contrastive features). Phonological rules then add allophonic details, capturing broader generalizations which apply to sounds in particular contexts across word forms. Finally, phonetic implementation rules determine the fine-grained phonetic details of how sounds ought to be produced in these contexts. Rules affecting a given segment in a particular context apply to all instances of that segment across the lexicon; because of this, there can be no synchronic gradient variation between words per se. In the abstractionist view, a word's surface realization is essentially the predictable phonetic sum of its parts.

This state of affairs also has diachronic implications: gradual phonetic shift that affects some words and not others on a lexically unpredictable basis should not occur. Phonetically speaking, words should not 'have their own history' (pace Malkiel 1967), but undergo regular, Neogrammarian sound change.

### *1.2.2 Phonetically rich view of representation*

A more recent view of representation is that the stored phonological knowledge of a particular word consists not of a minimal, abstract sequence of symbolic elements, but a large collection of phonetically rich memories of particular tokens of that word. This view characterizes the approach of usage-based theories such as those proposed by

[5] Scholars disagree on which features may be considered redundant and how they are filled in at later points in the phonological derivation (see e.g. Archangeli 1988; Steriade 1995). However, the details of underspecification theory are not important to the matter at hand; what matters here is the abstractness of these representations relative to phonetic forms.

[6] Abstractionist views do not prohibit lexical items from having multiple underlying representations. However, the positing of multiple representations is generally reserved for cases of lexically idiosyncratic phonemic variation. For example, variation in a word like *vase* can be captured by positing two underlying representations – /ves/ and /vaz/ – which then compete in some way for selection.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Bybee (2001) and developed by scholars working within Exemplar Theory (Johnson 1997; Pierrehumbert 2001, 2002, 2003; Wedel 2004, 2006), and has precursors in the Memory Trace Models described by e.g. Hintzman (1986) and Goldinger (1998), and the 'memory images' posited by Paul (1880). Again, it is helpful to tease apart various related components of this view.

In usage-based theories, the mental representation of lexical items reflects much of the phonetic detail of actual surface forms. In fact, they are often considered to be memories of utterances embedded within the parametric phonetic space, a 'quantitative map of the acoustic and articulatory space' (Pierrehumbert 2003: 179). Categories such as words, phonemes, and allophones are abstractions over this phonetic space: word forms correspond to clouds of remembered tokens associated with a given semantic label (e.g. DOG), and sound categories such as phonemes and allophones emerge as distributional peaks within this phonetic space which may receive their own labels (e.g. 'p', '$p^h$'). This proposal was initially motivated by experimental findings indicating that listeners retain memories of words spoken in particular voices (Hintzman *et al.* 1972; Cole *et al.* 1974; Mullennix *et al.* 1988) and with particular intonational contours (Schacter & Church 1992; Church & Schacter 1994), and will even adjust their perception of phonemes as a result of exposure to talker idiosyncrasies (e.g. Nygaard & Pisoni 1998; Norris *et al*. 2003). It has since been developed to account for linguistic phenomena such as lexically specific (often frequency-related) phonetic change (Phillips 1984) and diachronic phonetic shifts (see Pierrehumbert 2003).

Because each heard token of a lexical item is stored and tagged with a label indexing it to that lexical item, there are potentially hundreds or thousands of representations associated with each word. Usage-based theories differ with respect to how many memories are retained, and for how long (recent versions of Exemplar Theory, for instance, contain a decay parameter which allows older exemplars to be forgotten over time, e.g. Pierrehumbert 2006). In most such theories, however, the number of representations will vary depending on how often tokens of the word are encountered (whether in the speech of others or in that of the speaker herself), with more frequent words having more stored memories.

A common characteristic of usage-based models is the lack of a clear distinction between representations and rules (Langacker 1987, 2000; Bybee 2001). Phonological generalizations are not formalized as processes that representations undergo, but as emergent from the distributional regularities present across lexical representations. Another way to state this is to say that there is no derivationally based distinction between phonemes (qua the components of underlying representation) and allophones (qua the results of phonological rules). Both types of categories are represented by the distributional peaks which form among clouds of exemplars in the parametric phonetic space, with clouds corresponding to classical allophones being more circumscribed within this space than those corresponding to the higher-level classical phoneme.

Finally, in usage-based models, every word does in fact have its own history, reflecting the assumption that lexical representations are dynamic and affected by usage. Representations are continually updated with new heard tokens, but this process

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

varies across lexical items, such that frequently heard items will be updated more often than rarely heard items. Precise predictions regarding the effect of lexical frequency on sound change are difficult to nail down. As Pierrehumbert (2006) notes, the relative salience of certain items, saturation effects for the memory of high-frequency items, and other cognitive factors may mediate frequency effects. Moreover, while it seems intuitive that more frequently updated items will be more advanced with respect to change, it is also the case that the representations of frequently encountered items will contain many older exemplars, and the presence of this phonetic baggage might be expected to slow the progress of change.

### *1.2.3 The representation of contrast*

Contrast is represented rather differently in each of these views. In abstractionist theories, stating that two sounds contrast means that the segments '[differ] in at least one feature' (see Chomsky & Halle 1968: 336 for a formal definition). Thus contrast in this framework is a clearly binary notion, such that the phonological representations of two sounds/words either contrast (because they differ in at least one feature) or are identical; in this approach, there is no such thing as 'a small difference of sound' (Bloomfield 1926).

The findings on near-merger described above demonstrate that there is variation with respect to how differently two categories may be realized in actual speech. The words *cot* and *caught*, for example, may be realized with a large enough phonetic distance between them that no speaker could fail to detect the difference, or they may overlap in phonetic realization to such an extent that most speakers no longer remark upon the difference even though a small one continues to be made (of course, intermediate cases are also possible). All possibilities along this continuum, however, are represented in the same way in the abstractionist view: there are simply two distinct underlying representations, and the ultimate distance between realizations of these are determined by phonetic implementation rules operating on each category. Near-mergers thus occur when phonetic implementation rules realize two categories so similarly that speakers can no longer perceive the difference. Importantly, such near-merger effects are expected to apply across the lexical board: because any rule-based phonetic merging applies to all words containing the relevant category, this account implies that all words should participate in near-merger phenomena to the same extent.

Usage-based approaches do not draw such a firm line between the existence of contrast and its phonetic realization. At a certain level, contrast may also be considered a binary notion in such frameworks: either two clouds of tokens are associated with two different category labels (e.g. 'ɑ' and 'ɔ'), or the same category label. However, such labels do not exist prior to phonetic realizations, but emerge from instantiations of particular items which clump together in the phonetic space; gradience is thus built into these underlying representations. The clumps corresponding to particular category labels may be largely separate, or may overlap to varying extents. If two such clumps overlap to a great enough degree that speakers cannot reliably apply the right category label based on phonetic differences, then near-merger behavior may occur.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Contrast in such theories is not just phonetically gradient, but lexically gradient as well: the relevant bits of different words containing the 'same' vowel category may occupy somewhat different places in the parametric phonetic space, depending on the input a speaker gets for individual items. This model thus predicts that individual words (or, more to the point, potentially homophonous word pairs) may show greater or lesser contrast (e.g. *taught–tot* might show less separation than *caught–cot*).

Because they represent contrast in different ways, these two views also make different predictions regarding how new contrasts may be acquired. The next section describes each of these sets of predictions for acquisition of the low back vowel distinction, laying the foundation for the study described in section 2.

*1.2.4 Acquiring a new contrast in abstractionist models*

There is little in the generative phonology literature that addresses the issue of intraspeaker linguistic change beyond the age of L1 acquisition. However, we can speculate about the possibilities for intraspeaker change in an abstractionist framework based on the types of representations that would be changing.

To start, speakers who do not have a low back vowel contrast are assumed to store identical featural representations for lexical items such as *cot* /kɑt/ and *caught* /kɑt/. In order for 'complete' unmerging – in the sense of replication of a two-phoneme speaker's low back vowel output – to occur, every low back vowel in the one-phoneme speaker's lexicon must be altered to include an additional feature that will enable later rules (ultimately, the phonetic component) to realize the contrast. Such comprehensive acquisition of the contrast as realized in a low-back-vowel-distinguishing dialect seems unlikely, as the would-be two-phoneme individual may simply not be exposed to tokens of every low back vowel word in the new dialect. The unlikelihood of complete unmerging in this sense has been put forth as an argument for why mergers are necessarily irreversible, and as a explanation of Herzog's Principle (Labov 1994). However, this is a straw man; there is obviously a (logically possible) middle ground between learning a new sound for all relevant words and learning the sound for none of those words. If features can be added to underlying representations, then we might expect that these additions would occur on a word-by-word basis, with perhaps highly frequent and/or highly salient words acquiring a value for the new feature first. While this change would occur in a lexically gradual manner (in what may be termed a 'split-by-transfer', in parallel with the phenomenon of merger-by-transfer (Trudgill and Foxcroft 1978)), the results of it ought to be phonetically abrupt.[7] That is, words may vary in terms of when they receive their new feature value, but because the words will be receiving one of two values for that new feature, they should ultimately be spelled-out in one of two ways: any word that has received a new feature value as a result of the split should be realized in essentially the same way as every other word that has received that same new value. The magnitude of the phonetic distance between

[7] 'Phonetically abrupt' is a bit of a Neogrammarian misnomer for the analogical replacement of one phoneme with another. I use the usual terminology here.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

these two spell-outs is difficult to predict; it might be small or large, depending on the nature of the input a speaker receives.

*1.2.5 Acquiring a new contrast in phonetically rich theories*

Usage-based phonology has a much more clearly defined account of intraspeaker change. This is, of course, because dynamic phonological representations are at the core of this type of theory. In the view discussed in the previous section, the underlying representation of a word is abstract and mostly fixed, with later rules left to do most of the heavy lifting in terms of variation and change. However, in a usage-based model, the word-level representation is the primary locus of change: new tokens of words cause shifts in the phonetic distribution of their associated exemplar clouds, and changes at the level of phonological categories (which comprises generalizations over these word forms) follow from these changes in distributional weightings.

Acquiring a new contrast is thus predicted to occur in a very different manner from that described in the previous section. In this case, the one-phoneme speaker starts out with two lexical items, *cot* and *caught*, each of which is associated with a cloud of exemplars. Unlike those of the two-phoneme speaker, these clouds are largely coterminous in the phonetic space. If the one-phoneme speaker is exposed to a dialect in which these words are realized differently (with, for example, tokens of *caught* occupying a higher and backer region of the parametric phonetic space than tokens of *cot*), the phonetic distributions of their associated clouds will gradually diverge.[8]

As noted above, precise frequency predictions regarding the way in which splits should be acquired are difficult to make. Setting aside the mediating effects of cognitive factors such as word salience, it is not clear how high-frequency items should pattern in an unrefined usage-based model in which all tokens are retained and given equal weight: the frequent accrual of new tokens may result in a high-frequency item being more advanced with respect to a change, but the same item's large collection of old tokens may serve to slow its progress. However, in a model in which older exemplars are assumed to decay and newer items can have more influence (Pierrehumbert 2001), the predictions are clearer: high-frequency items should show signs of change before low-frequency items. Moreover, this change should be phonetically gradual: words do not receive one of two feature values which divide them into two phonetic groups, but instead are expected to shift gradually in the phonetic space, reflecting the ongoing incorporation of gradiently variable heard tokens into representational clouds laden with older remembered exemplars.

[8] Misunderstandings may occur, with resulting occasional mis-storages of tokens. Labov (2010) discusses the frequency and relevance of natural misunderstandings as a result of dialect change, noting that 14 percent of the misunderstandings in his corpus are tied to the low back vowel merger. Most of these, however, seem to implicate the pairs *Don–Dawn* (names) and *copy–coffee* (nouns), whose members may occupy the same syntactic position in an utterance. It is harder to imagine many cases where *cot* (a noun) could be confused with *caught* (a verb). In any case, it seems unlikely that such misunderstandings would have a great systematic effect on the representations of relevant words.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

We thus have two different sets of predictions regarding how new contrasts should be acquired. In both views of representation described here, we might expect a split to manifest itself first in words which are more often encountered. In the phonetically rich view, this split is expected to be phonetically gradual, with more frequent items showing incrementally more advanced phonetic shift in the direction of the ambient dialect. In the abstractionist view, this split should be phonetically abrupt, reflecting a categorical change in the underlying representation for a word.

It is in principle possible to test these predictions by observing the behavior of speakers who are part of a community undergoing a split-in-progress, and determining whether these speakers show evidence of lexically gradual and phonetically gradual shift. As noted previously, however, splits-in-progress at the community level are rare compared to mergers-in-progress, so finding relevant data sets can be difficult. An alternative approach is to find native speakers of a dialect characterized by some merger who have been exposed to new dialect input which does not have this merger.

## 2 The study: contrast acquisition by Canadians in the New York region

The *Atlas of North American English* (ANAE) reports Canada to be a region characterized by merger of the (o) word class (encompassing *cot* and other words descended from the Middle English 'short-o' class) and the (oh) word class (including *caught* and other words mostly from the Middle English 'au' class) (Labov *et al.* 2006). According to Boberg (2008),'virtually all native speakers of Canada today' have this merger, which has been present in Canadian English for several generations. The situation in New York City and surrounding areas is quite different: this region is noted in ANAE as being one of a few areas in which the low back vowel distinction remains robust, with the raised quality of the vowel in (oh) words like *caught* being a particularly salient feature of the local dialect.

A person who acquires their native variety of English in Canada will start out with one low back vowel category, such that words in the (o) and (oh) word classes will not be distinguished in vowel quality. If such a person moves to the New York region, they will be exposed to dialect input in which (o) and (oh) words are realized with different qualities. A study of Canadians who have moved to the New York City region thus provides an opportunity to observe how speakers may go about acquiring a new contrast over time and to test the predictions made by the abstractionist and phonetically rich views of representation outlined above.

Such a project also allows us to approach the methodological questions of section 1.1 from a new angle. Studies of low back merger in progress have shown that the norms reflected in minimal pair tests – a speaker's knowledge *that* two sounds are the same or different – may not match up with *how* that speaker generally produces relevant words. Presumably the same kind of mismatches might characterize the behavior of speakers acquiring a split, but these have yet to be empirically established.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

### *2.1 Methods*

Sociolinguistic interviews were conducted in New York City and neighboring counties in New Jersey with 17 native Canadians who had moved to the New York metropolitan region as adults (after the age of 21). All interviews were recorded directly to 16 bit, 44.1Hz WAV files using an Edirol (by Roland) R-09 digital recorder and an Audio-Technica electret condenser lapel mic. Each interview was about an hour and a half long. Interviews began with basic questions about the speaker's background and where they grew up in Canada, later moving to their reasons for coming to the United States and their experience doing so. Speakers were asked for opinions of the area where they grew up and their adopted region, and were encouraged to compare their new and old homes at both a local and national level (e.g. Toronto vs. New York City, Canada vs. the US). After about an hour of *conversation*, each speaker completed a *word list* reading, *minimal pair & rhyming tasks* and an *other dialect judgment task.* After these tasks, the conversation resumed with discussion of language and accent issues.

#### *2.1.1 Word list readings*

Speakers were asked to read out loud 135 words which were presented on flashcards. These items represented a variety of word classes, though (o) and (oh) words featured prominently in the list. Many of these low back vowel words were also present in the minimal pair list, enabling a comparison of vowel production across styles. Two versions of this word list were used over the course of data collection. The original word list (presented to the first five speakers interviewed) included fewer low back vowel words; once it became apparent that there were differences in how these vowels were produced across word list and minimal pair styles, more of the minimal pair list words were added to the word list to enable a more robust comparison across contexts for the remaining twelve speakers.

#### *2.1.2 Minimal pair/rhyming tasks*

Speakers also completed a sociolinguistic minimal pair task and a rhyming pair task. Each speaker was handed a printed list of minimal pairs, and asked to read each pair out loud, then say whether the pair sounded the same or different. Speakers were also given a shorter list of rhyming pairs and asked to pronounce each pair, then say whether the pair rhymed. Each of these lists primarily probed the low back vowel distinction, though these pairs were interspersed with other pairs of potential interest (e.g. *Mary–merry*).

#### *2.1.3 Other dialect judgment task*

After completing the canonical minimal and rhyming pair task, speakers were then asked to look back over these two lists and say whether they thought people from the New York region would either have different judgments of some of these pairs, or pronounce particular words differently. The purpose of this task was to determine whether speakers are aware of the low back vowel distinction in New York-area English. When speakers identified specific pairs as being produced differently in the local dialect, they were encouraged to produce these forms 'as a local would say them', so that I

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

might get a better sense of what they believed the local phonetic targets for relevant words to be.

Tokens of low back vowels from each of these four contexts – *conversation*, *word list*, *minimal pair/rhyming* and *other dialect judgment* – were acoustically and, where appropriate, statistically analyzed to answer the following questions:

- Is there a phonetic difference between (o) words and (oh) words in any context, and if so, what is the magnitude of this difference?
- In cases where a split seems to have occurred, has this happened in a lexically gradual manner?
- Are speakers aware of the (o)/(oh) contrast, either in their own speech or in the ambient dialect?
- Is there any relationship between awareness of the contrast (either in one's own speech or in that of local dialect speakers) and production of the contrast?

### *2.2 Acoustic and statistical analysis*

Measurements of F1 and F2 were taken for each low back vowel token at the F1 maximum, the point representing the lowest point of the vowel. Measurement points were first marked automatically with a script in Praat, then manually checked for egregious errors and, if necessary, corrected. Vowel duration was also measured in the minimal pair and word list contexts.

To determine whether each speaker produced a distinction between (o) words and (oh) words in the minimal pair and word list contexts, F1, F2, and duration was compared across the two word classes in each context using paired t-tests.[9]

To determine whether a distinction is made in conversational speech, every useable token of words from the (o) and (oh) classes was extracted from the portion of each speaker's recorded interview that took place before the reading and judgment tasks. 'Useable' in this case means any token that showed reasonable formant tracking in Praat; tokens produced with excessively creaky or falsetto voice quality, or against background noise, were excluded. Auditorily reduced tokens were also excluded; in practice this meant any vowel with a duration of less than 50 milliseconds. All selected tokens had primary or secondary stress on the low back vowel. Tokens were classified as either (o) or (oh) based on how each word is produced in the New York/New Jersey varieties of English which make this distinction. Across all 17 speakers, 2,736 conversational tokens of (o) words and 1,487 tokens of (oh) words were collected for measurement. Each token was coded for *word class* (o or oh) and four phonological context factors: *preceding place*, *following place*, *preceding voice/manner* and *following voice/manner.*

[9] Paired t-tests are ideal for cases in which the between-group variation is small compared to the variation within those groups. This, of course, is exactly the situation faced in determining whether the Canadian speakers in this study are producing a (o)/(oh) distinction in their minimal pairs: the difference between the two word classes is likely to be slight, while the differences across pairs due to varying phonological contexts is likely to be great. Using the more powerful paired t-test increases the likelihood that any difference between (o) and (oh) words in this list will be detected.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

The analysis of the conversational data required more than a simple comparison of mean measurement values, for two major reasons. First, unlike those elicited in minimal pair tasks and carefully constructed word lists, vowel tokens plucked from natural conversation are not balanced in terms of phonological environment. This is an especially relevant concern for the (o) and (oh) word classes, which are distributed unevenly across phonological contexts for reasons having to do with the historical development of these classes (Labov *et al.* 2006). It is thus necessary to account for the effects of phonological context in the analysis to ensure that acoustic differences arising from different contexts are not mistaken for phonologically unpredictable variation. Second, given that every useable token of a relevant word was included in the analysis, it is desirable to have a way of factoring in possible word-specific effects, to ensure that particular overrepresented words in the sample do not skew the results.

To address both of these issues, mixed effects regression analysis was implemented using the lmer() function in R (Bates & Sarkar 2008; Pinheiro & Bates 2000; Baayen 2008). For each formant, for each speaker, a model was created that included fixed effects corresponding to the four phonological context variables described above, a fixed effect of word class (o vs. oh) and a random effect of word. This model was compared with a simpler model containing the same fixed phonological effects and the random effect of word but no word class term, to determine whether adding word class results in a significantly better model.

Two pieces of information result from this procedure. First, the comparison of the two models revealed whether a speaker exhibits low back vowel variation which is at least partially predicted by word class membership after phonological context has been taken into account – that is, whether there is evidence of low back vowel contrast in that speaker's conversational speech. Second, the effect size associated with word class in the more complex model can be interpreted as a measure of the distance in Hz between (o) and (oh), also after the effects of phonological context have been taken into account.

## *2.3 Results*

### *2.3.1 Minimal pairs*

For each speaker, the minimal pair/rhyming task yields two results: a perception result (whether they perceive a difference in their own speech) and a production result (whether these two word classes are produced distinctly).

All speakers uniformly reported that the (oh)/(o) pairs sounded 'the same' after producing them, thus exhibiting a 'merger in perception' with respect to these two word classes.

Nearly all speakers were also merged in production (see tables 1–3). No significant difference was found for any measure between the two vowels in this style, with one exception: JC's mean (oh) F2 is 31Hz lower than his mean (o) F2 ($t(9) = -2.6664$, $p = 0.03$), indicating a slight difference in backing consistent with how these word classes are realized in New York. This single significant result may very well be a chance

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Table 1. *Minimal pair test production results: F1 (means and standard deviations in Hz)*

| Speaker | (o) F1 Mean | SD | (oh) F1 Mean | SD | Mean difference | t(df) | p |
|---|---|---|---|---|---|---|---|
| BK | 746 | 98 | 727 | 125 | 19 | $t(8) = -0.5432$ | 0.60 |
| BW | 624 | 25 | 626 | 30 | −2 | $t(9) = 0.2839$ | 0.78 |
| CW | 805 | 27 | 793 | 58 | 12 | $t(8) = -0.7821$ | 0.46 |
| DB | 685 | 46 | 696 | 67 | −11 | $t(9) = 0.8040$ | 0.44 |
| ES | 614 | 58 | 595 | 79 | 19 | $t(9) = -0.7242$ | 0.49 |
| EW | 612 | 31 | 608 | 58 | 4 | $t(9) = -0.2481$ | 0.81 |
| GH | 713 | 68 | 708 | 75 | 5 | $t(9) = -0.4123$ | 0.69 |
| JC | 652 | 61 | 645 | 58 | 7 | $t(9) = -0.8615$ | 0.41 |
| JF | 684 | 47 | 693 | 47 | −9 | $t(9) = 1.0446$ | 0.32 |
| LC | 782 | 66 | 793 | 90 | −11 | $t(9) = 0.6808$ | 0.51 |
| LG | 722 | 80 | 746 | 81 | −24 | $t(9) = 1.0737$ | 0.31 |
| LW | 758 | 42 | 764 | 91 | −6 | $t(9) = 0.1596$ | 0.88 |
| NW | 745 | 114 | 744 | 110 | 1 | $t(9) = -0.0403$ | 0.97 |
| PW | 669 | 57 | 654 | 46 | 15 | $t(9) = -1.1469$ | 0.28 |
| SS | 670 | 72 | 660 | 77 | 10 | $t(8) = -0.6105$ | 0.56 |
| TM | 766 | 55 | 737 | 63 | 29 | $t(9) = -1.3920$ | 0.20 |
| VJ | 661 | 74 | 656 | 145 | 5 | $t(8) = -0.1125$ | 0.91 |

Table 2. *Minimal pair test production results: F2 (means and standard deviations in Hz)*

| Speaker | (o) F2 Mean | SD | (oh) F2 Mean | SD | Mean difference | t(df) | p |
|---|---|---|---|---|---|---|---|
| BK | 1127 | 205 | 1124 | 184 | 3 | $t(8) = -0.0345$ | 0.97 |
| BW | 1010 | 63 | 1003 | 60 | 7 | $t(9) = -0.5399$ | 0.60 |
| CW | 1149 | 49 | 1116 | 61 | 33 | $t(8) = -1.3901$ | 0.20 |
| DB | 1019 | 76 | 1032 | 93 | −13 | $t(9) = 0.7717$ | 0.46 |
| ES | 1055 | 132 | 1030 | 88 | 25 | $t(9) = -0.8064$ | 0.44 |
| EW | 924 | 60 | 929 | 83 | −5 | $t(9) = 0.2381$ | 0.82 |
| GH | 1095 | 95 | 1086 | 110 | 9 | $t(9) = -0.6101$ | 0.56 |
| JC | 984 | 79 | 952 | 82 | 32 | $t(9) = -2.6664$ | 0.03 |
| JF | 1022 | 64 | 1040 | 52 | −18 | $t(9) = 1.0293$ | 0.33 |
| LC | 1014 | 58 | 1078 | 157 | −64 | $t(9) = 1.2168$ | 0.25 |
| LG | 1017 | 51 | 1009 | 100 | 8 | $t(9) = -0.2761$ | 0.79 |
| LW | 1142 | 97 | 1121 | 107 | 21 | $t(9) = -0.5767$ | 0.58 |
| NW | 1180 | 77 | 1181 | 90 | −1 | $t(9) = 0.0294$ | 0.98 |
| PW | 1048 | 107 | 1009 | 61 | 38 | $t(9) = -2.0037$ | 0.08 |
| SS | 1101 | 49 | 1087 | 72 | 14 | $t(8) = -0.9279$ | 0.38 |
| TM | 1298 | 169 | 1223 | 154 | 75 | $t(9) = -1.8134$ | 0.10 |
| VJ | 1098 | 78 | 1117 | 65 | −19 | $t(8) = 0.6641$ | 0.53 |

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Table 3. *Minimal pair test production results: duration (means and standard deviations in ms)*

| Speaker | (o)duration Mean | SD | (oh)duration Mean | SD | Mean difference | t(df) | p |
|---|---|---|---|---|---|---|---|
| BK | 164 | 42 | 169 | 64 | −5 | t(8) = 0.3636 | 0.73 |
| BW | 212 | 50 | 214 | 47 | −2 | t(9) = 0.1632 | 0.87 |
| CW | 235 | 59 | 263 | 87 | −28 | t(8) = 1.0067 | 0.34 |
| DB | 263 | 70 | 270 | 58 | −7 | t(9) = 0.5324 | 0.61 |
| ES | 234 | 74 | 236 | 84 | −2 | t(9) = 0.2073 | 0.84 |
| EW | 214 | 46 | 215 | 61 | −1 | t(9) = 0.0658 | 0.95 |
| GH | 211 | 57 | 218 | 61 | −7 | t(9) = 0.6543 | 0.53 |
| JC | 210 | 73 | 216 | 90 | −6 | t(9) = 0.4566 | 0.66 |
| JF | 187 | 75 | 190 | 71 | −3 | t(9) = 0.2244 | 0.83 |
| LC | 244 | 70 | 257 | 89 | −13 | t(9) = 1.3847 | 0.20 |
| LG | 198 | 64 | 202 | 60 | −4 | t(9) = 0.4251 | 0.68 |
| LW | 248 | 65 | 264 | 84 | −16 | t(9) = 1.0386 | 0.33 |
| NW | 235 | 79 | 236 | 75 | −1 | t(9) = 0.0613 | 0.95 |
| PW | 233 | 78 | 241 | 87 | −8 | t(9) = 0.4961 | 0.63 |
| SS | 264 | 98 | 286 | 113 | −22 | t(8) = 2.1928 | 0.06 |
| TM | 210 | 55 | 235 | 82 | −25 | t(9) = 1.5736 | 0.15 |
| VJ | 209 | 71 | 237 | 138 | −28 | t(8) = 0.6088 | 0.56 |

occurrence. However, it may also be grounded in the particular linguistic history of this speaker, whose father was born in Brooklyn.

Aside from JC, however, 16 of the 17 speakers show a merger in production consistent with their merger in perception. In this style, at least, they do not seem to be showing much accommodation towards the New York-area contrast, instead patterning like native speakers of Canadian English.

### *2.3.2 Word lists*

More complicated results emerge from the word list data. The original point of the word list in this study was simply to elicit a few tokens of every lexical set, with the aim of establishing a citation form vowel space. Thus the first version of the word list, administered to the first five speakers interviewed, contained just 7 (oh) words and 5 (o) words. However, it became apparent that speakers were producing these words differently across the two read styles: for speakers BK, GH, JC, SS and VJ, (o) and (oh) words were auditorily more distinct in word list style, and showed greater separation in the vowel space (see figures 1–5). Though a significant difference between (o) and (oh) in either dimension could not be established given the small number of tokens for these speakers, these impressionistic results indicated the need for a more deliberate investigation of the low back vowel contrast in word list versus minimal pair style.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

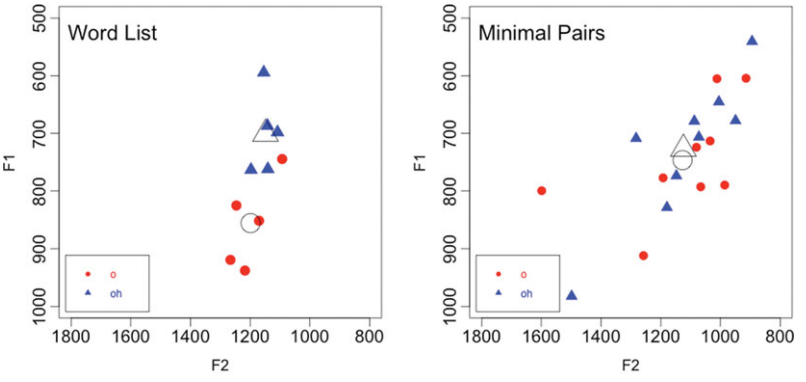



Figure 1. (Colour online) BK's low back vowel productions in read styles

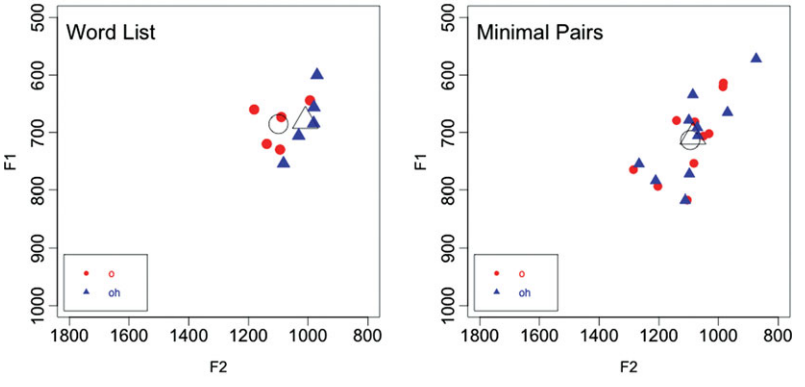



Figure 2. (Colour online) GH's low back vowel productions in read styles

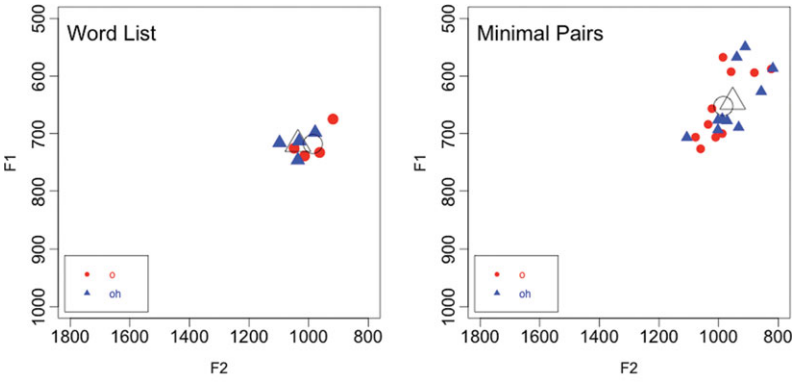



Figure 3. (Colour online) JC's low back vowel productions in read styles

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

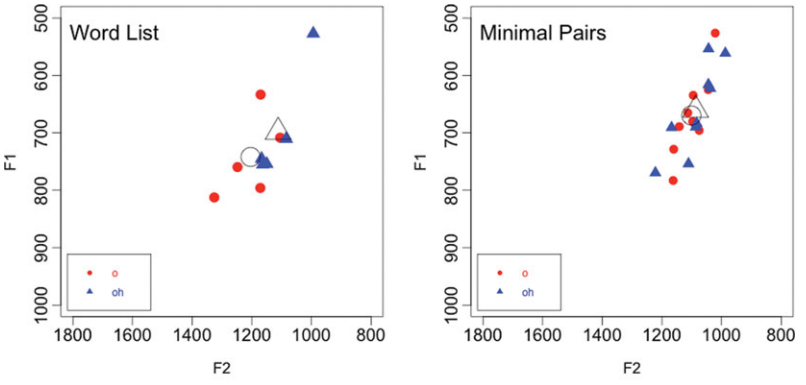



Figure 4. (Colour online) SS's low back vowel productions in read styles

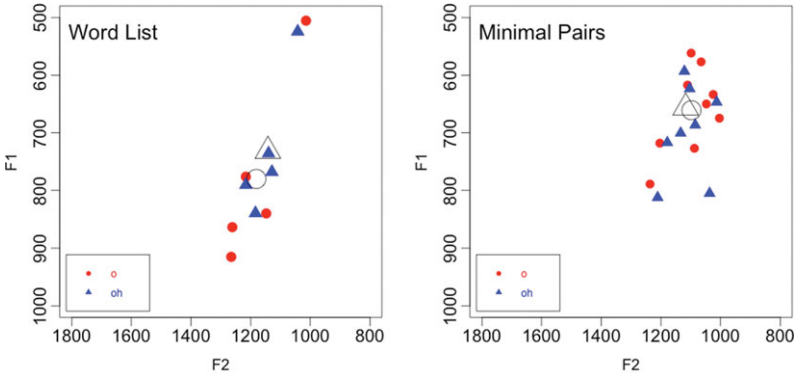



Figure 5. (Colour online) VJ's low back vowel productions in read styles

The word list was accordingly expanded to include all of the low back vowel word pairs already included in the minimal pair list. This change enabled a statistical examination of whether a contrast was present in the word list style alone, as well as a comparison of words across styles to see whether a shift had taken place in one or both vowels.

Several patterns of results were found among the group of twelve speakers who read the second version of the word list; these results are listed in tables 4–6.

For BW, DB, EW, LC and JF, no significant difference was detected in any dimension between (o) and (oh) in word list style. There also appears to be no appreciable shift in vowel quality across read styles. For these speakers, the two word classes occupy essentially the same vowel space in both word list and minimal pair context (figures 6–10).

ES and LW showed no significant difference between word classes in either formant measure, though ES's (oh) is significantly longer than his (o) in word list style. While these speakers do not seem to distinguish two vowels in either minimal pair or word list productions, there is some indication of a change in vowel quality across these tasks:

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Table 4. *Word list production results: F1 (all means and standard deviations in Hz)*

| Speaker | (o) F1 Mean | SD | (oh) F1 Mean | SD | Mean difference | t(df) | p |
|---|---|---|---|---|---|---|---|
| BW | 629 | 28 | 636 | 21 | −7 | $t(14) = 1.1928$ | 0.25 |
| CW | 848 | 45 | 809 | 63 | 39 | $t(13) = -2.3821$ | 0.03 |
| DB | 673 | 73 | 659 | 109 | 14 | $t(13) = -0.5023$ | 0.62 |
| ES | 665 | 85 | 662 | 55 | 3 | $t(13) = -0.1493$ | 0.88 |
| EW | 589 | 33 | 579 | 35 | 10 | $t(14) = -1.0650$ | 0.30 |
| JF | 709 | 79 | 690 | 59 | 19 | $t(13) = -1.0650$ | 0.21 |
| LC | 791 | 54 | 782 | 76 | 9 | $t(13) = -0.5638$ | 0.58 |
| LG | 750 | 123 | 687 | 137 | 63 | $t(14) = -1.5984$ | 0.13 |
| LW | 848 | 70 | 825 | 72 | 23 | $t(13) = -0.8265$ | 0.42 |
| NW | 835 | 91 | 757 | 110 | 78 | $t(13) = -2.6927$ | 0.02 |
| PW | 681 | 46 | 662 | 46 | 19 | $t(14) = -1.6748$ | 0.12 |
| TM | 778 | 99 | 727 | 84 | 51 | $t(13) = -1.8892$ | 0.08 |

Table 5. *Word list production results: F2 (all means and standard deviations in Hz)*

| Speaker | (o) F2 Mean | SD | (oh) F2 Mean | SD | Mean difference | t(df) | p |
|---|---|---|---|---|---|---|---|
| BW | 1058 | 65 | 1057 | 47 | 1 | $t(14) = -0.0804$ | 0.94 |
| CW | 1144 | 77 | 1112 | 130 | 32 | $t(13) = -0.9299$ | 0.37 |
| DB | 1053 | 62 | 1040 | 78 | 13 | $t(13) = -0.5709$ | 0.58 |
| ES | 1065 | 139 | 1078 | 54 | −13 | $t(13) = 0.4407$ | 0.67 |
| EW | 926 | 62 | 933 | 56 | −7 | $t(14) = 0.5638$ | 0.58 |
| JF | 1049 | 81 | 1026 | 74 | 23 | $t(13) = -1.6555$ | 0.12 |
| LC | 1042 | 81 | 1044 | 96 | −2 | $t(13) = 0.1341$ | 0.90 |
| LG | 1062 | 85 | 978 | 104 | 84 | $t(14) = -2.650$ | 0.02 |
| LW | 1215 | 99 | 1206 | 58 | 9 | $t(13) = -0.3287$ | 0.75 |
| NW | 1218 | 70 | 1182 | 79 | 36 | $t(13) = -1.6947$ | 0.11 |
| PW | 1047 | 77 | 1000 | 103 | 47 | $t(14) = -2.1575$ | 0.049 |
| TM | 1300 | 85 | 1258 | 84 | 42 | $t(13) = -2.4267$ | 0.03 |

their apparently single low back vowel is slightly fronter and lower in word list style than in minimal pair style (figures 11–12).

CW and NW both show a significant difference in F1 between (oh) and (o) in word list style. In both cases, it appears that (o) is lower in word list tokens than in minimal pairs; (oh), meanwhile, does not seem to vary much between contexts (figures 13–14). Finally, speakers LG, PW and TM show significant differences in F2 between (o) and (oh) in word list style. Again, this difference mainly seems to be due to variation in (o) across contexts, though LG's (oh) also appear to be somewhat backer in word list forms (figures 15–17).

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Table 6. *Word list production results: duration (all means and standard deviations in ms)*

| Speaker | (o)duration Mean | SD | (oh)duration Mean | SD | Mean difference | t(df) | p |
|---|---|---|---|---|---|---|---|
| BW | 262 | 66 | 250 | 60 | 12 | $t(14) = -1.0266$ | 0.32 |
| CW | 236 | 82 | 244 | 89 | −8 | $t(13) = 0.9361$ | 0.37 |
| DB | 225 | 62 | 231 | 77 | −6 | $t(13) = 0.4317$ | 0.67 |
| ES | 169 | 44 | 198 | 69 | −28 | $t(13) = 2.3137$ | 0.04 |
| EW | 183 | 34 | 191 | 47 | −8 | $t(14) = 1.3686$ | 0.19 |
| JF | 164 | 73 | 173 | 63 | −9 | $t(13) = 0.8233$ | 0.43 |
| LC | 210 | 72 | 206 | 71 | 4 | $t(13) = -0.2723$ | 0.79 |
| LG | 176 | 55 | 195 | 80 | −19 | $t(14) = 1.3296$ | 0.20 |
| LW | 166 | 76 | 161 | 59 | 5 | $t(13) = -0.4291$ | 0.67 |
| NW | 145 | 39 | 184 | 50 | −39 | $t(13) = 4.4024$ | <0.001 |
| PW | 188 | 66 | 182 | 80 | 6 | $t(14) = -0.4211$ | 0.68 |
| TM | 162 | 56 | 169 | 54 | −7 | $t(13) = 0.5812$ | 0.57 |

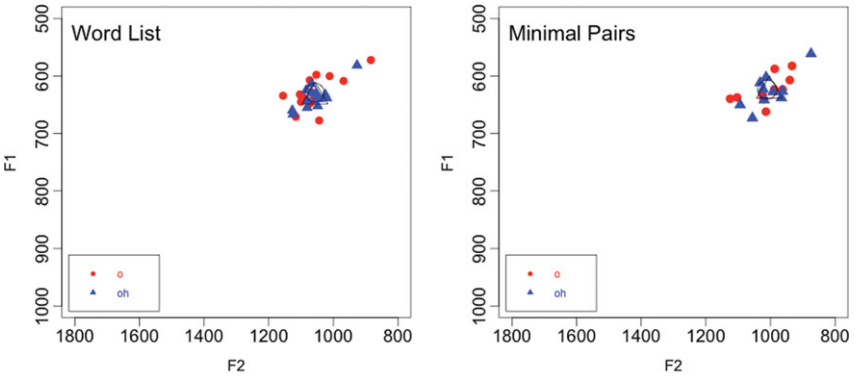



Figure 6. (Colour online) BW's low back vowel productions in read styles

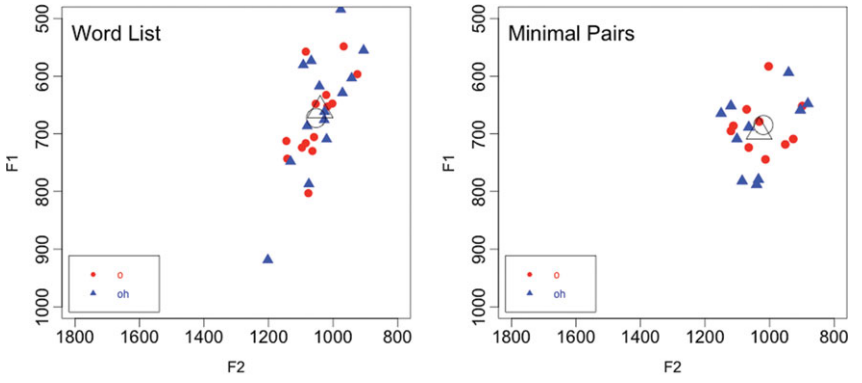



Figure 7. (Colour online) DB's low back vowel productions in read styles

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

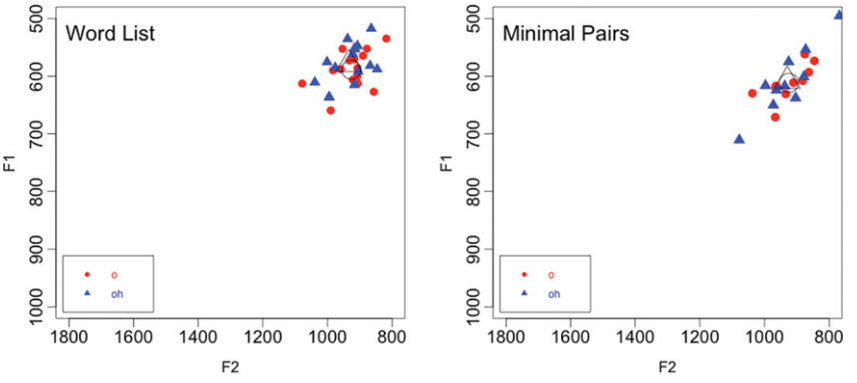



Figure 8. (Colour online) EW's low back vowel productions in read styles

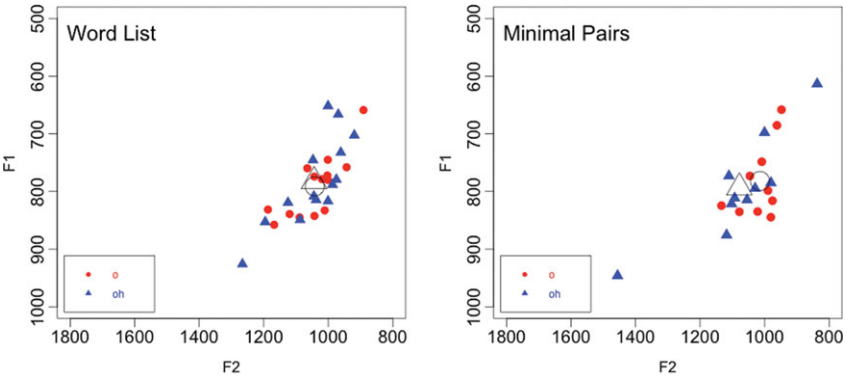



Figure 9. (Colour online) LC's low back vowel productions in read styles

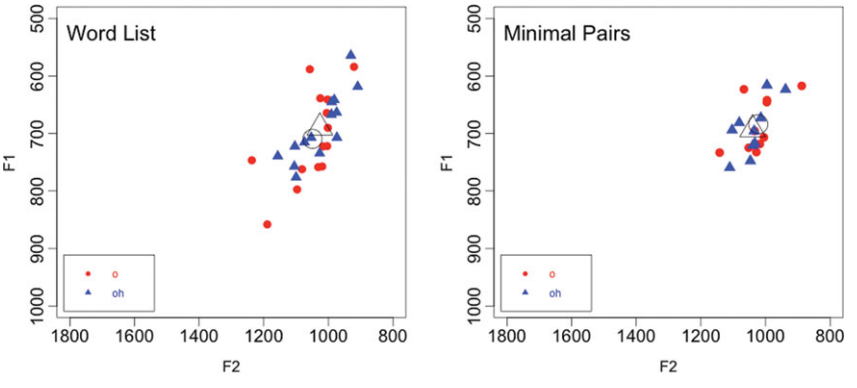



Figure 10. (Colour online) JF's low back vowel productions in read styles

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

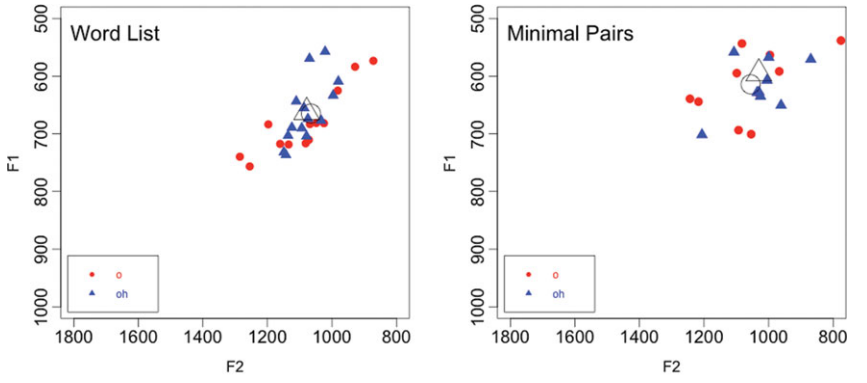



Figure 11. (Colour online) ES's low back vowel productions in read styles

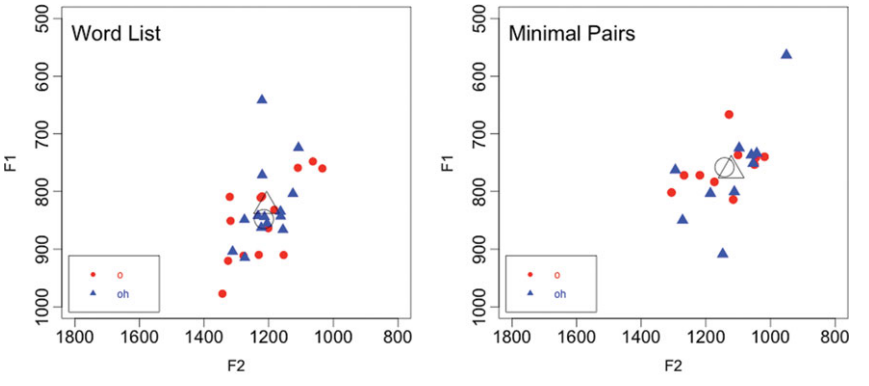



Figure 12. (Colour online) LW's low back vowel productions in read styles

To summarize, half of the 12 speakers who read the second, fuller word list distinguished between (o) and (oh) in this style along some phonetic dimension. A visual comparison of the vowel plots for each speaker indicates that those speakers who vary vowel quality across the two styles do so in a consistent manner. The speakers who produce a significant quality distinction in word list productions seem to be producing their (o) word class in a fronter and/or lower position (that is, closer to the realization of this word class for a speaker who has this distinction in the New York region). Meanwhile, even two speakers who did not distinguish (o) and (oh) in word list nonetheless produce their single undifferentiated vowel in a fronter and lower position.

### *2.3.3 Conversational data*

The results of the mixed effects analyses of conversational speech indicate that 11 of the 17 speakers produce a distinction between (o) words and (oh) words in some dimension in this context. These results are summarized graphically in figure 18, which plots the effect size (in Hz) associated with word class obtained in the F2 and F1 analysis

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

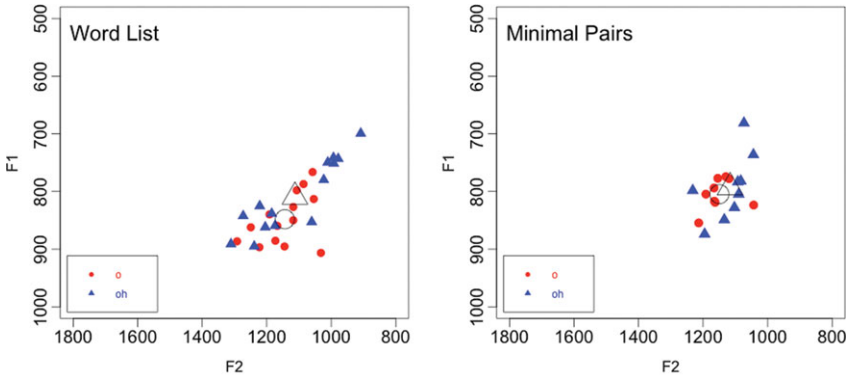



Figure 13. (Colour online) CW's low back vowel productions in read styles

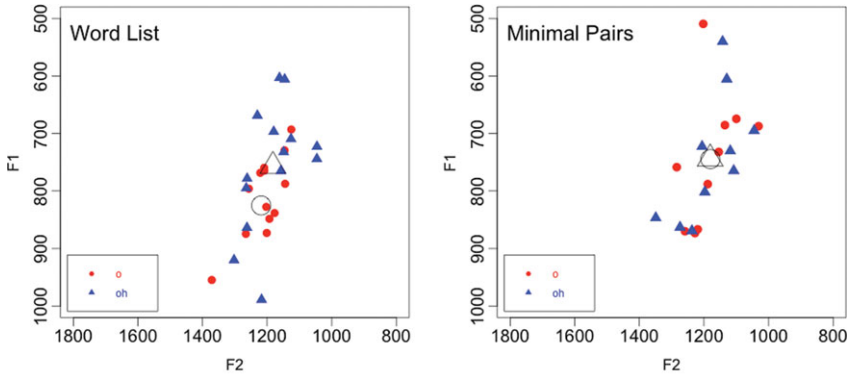



Figure 14. (Colour online) NW's low back vowel productions in read styles

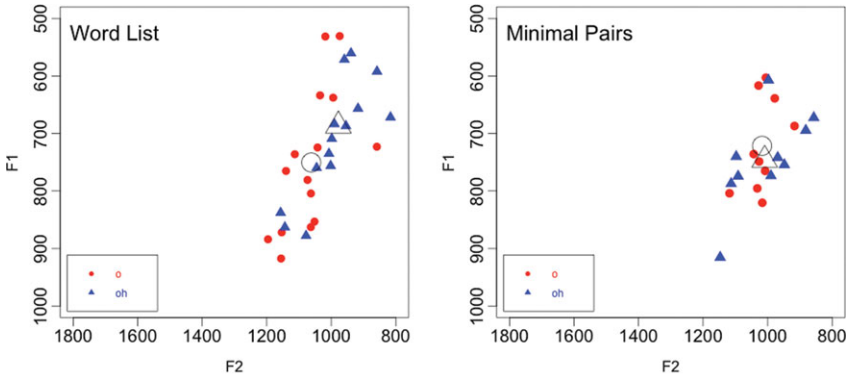



Figure 15. (Colour online) LG's low back vowel productions in read styles

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

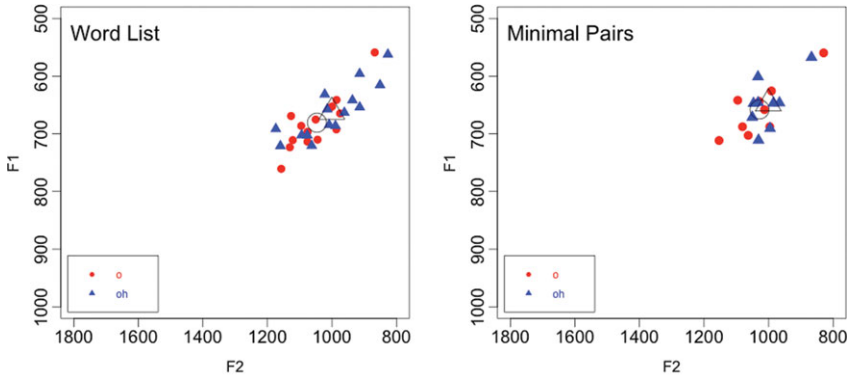



Figure 16. (Colour online) PW's low back vowel productions in read styles

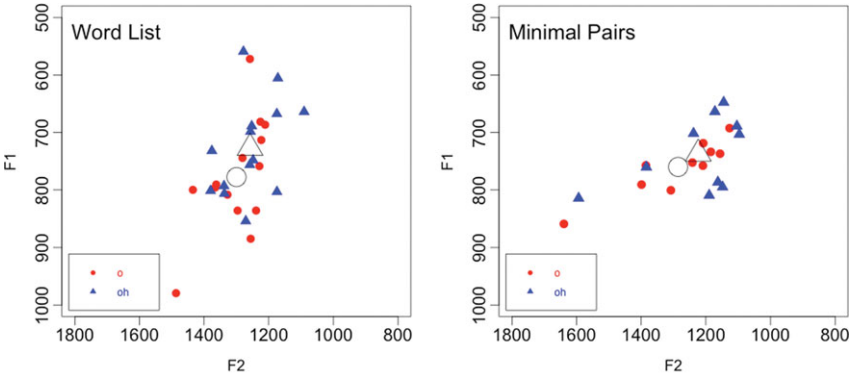



Figure 17. (Colour online) TM's low back vowel productions in read styles

of each speaker. Speakers with a large difference along both dimensions are plotted farther away from the origin, while speakers with very small effect sizes appear closer to the origin. Symbols surround the initials of those speakers for whom word class was found to be significant on one or both dimensions (upward pointing triangle = word class significant for F1 only; downward pointing triangle = word class significant for F2 only; diamond = significant for both formants).

A few points arise from these conversational results. First, while 11 speakers show a significant difference along at least one dimension in this context, there is wide variation in terms of how this difference is realized. SS, the speaker with the most robust distinction, has a Euclidean distance of 116Hz between (o) and (oh), while the distance between BW's (o) and (oh) words is only 38Hz. Second, even among speakers with no significant difference along either dimension, effects trend in the same direction. (o) words are associated with positive effects on both F1 and F2 – that is, (o) words are generally realized fronter and lower than (oh) words. For all speakers, however, there is still much phonetic overlap between (o) and (oh) words. Figure 19

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

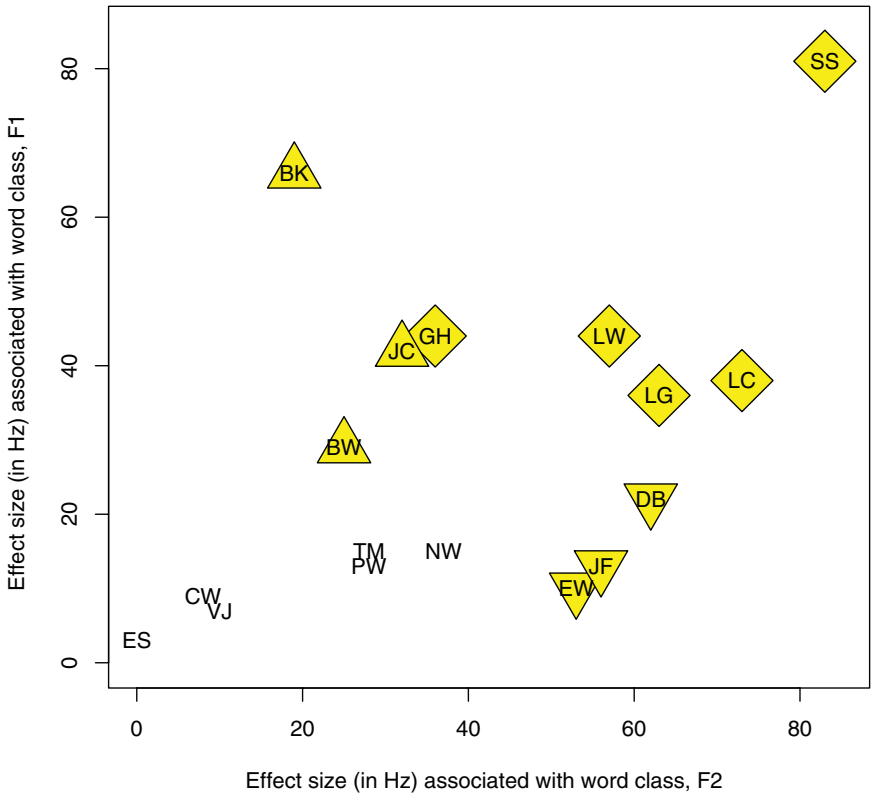



Figure 18. (Colour online) Results of the mixed effects analyses of conversational data. Speakers are plotted according to the effect sizes associated with word class for each of F1 and F2

contains scatterplots showing the distribution of tokens of both word classes in the conversational speech of the 5 speakers who make a significant distinction between these classes in both F1 and F2. Even for these 5 most distinct speakers, there is no clear separation between (o) and (oh).

It is also interesting to note the discrepancy in findings between word list and conversation, the two contexts in which some speakers make a distinction between (o) and (oh). That is, the set of speakers who distinguish these word classes in word list and the set of those who distinguish them in conversation are not identical, nor do they participate in any sensible subset relation.

*2.3.3.1 Frequency effects on contrast acquisition* The analysis of conversational data thus far has established that natively one-phoneme speakers may come to make a distinction between (o) and (oh) in spontaneous speech. In this section I will show that this distinction is also acquired in a lexically gradual manner, by demonstrating that there are frequency effects on the realization of (o) and (oh).

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

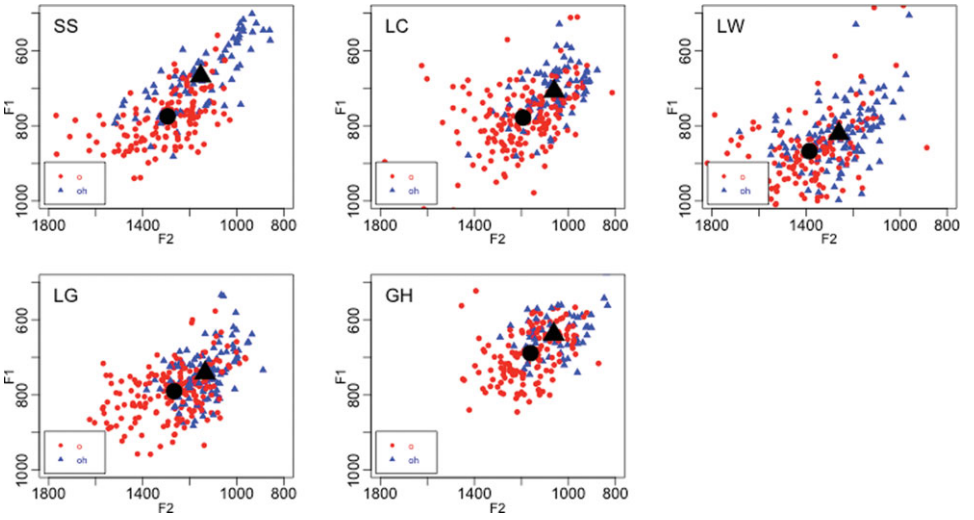



Figure 19. (Colour online) Scatterplots of conversational data for five speakers who distinguish (o) and (oh) in both F1 and F2

Two issues arise here. First, it is necessary to determine the right measure of frequency. Various corpora exist from which frequency counts can be obtained, but these fall short in various ways: many are based on written speech (e.g. CELEX (Baayen *et al.* 1993)), some are based on dialects of English which are not spoken by the speakers in this study (e.g. the British National Corpus) and others are simply out of date (e.g. the Brown Corpus (Kučera & Francis 1967)). Moreover, while certain words occur with high frequency in all 17 interviews, reflecting the commonality of these words in the linguistic input of all speakers, other words are idiosyncratically frequent, in ways which seem to reflect the individual lived experience – and likely linguistic input – of each speaker. For this reason, a speaker-internal measure of frequency was used. For example, the word *dog* is coded as frequency 6 for a speaker who uses that word six times, but as 2 for a speaker who uses it only twice over the course of an interview. Frequency counts here are simply raw counts of usage over the course of the interview. However, as all interviews were of roughly comparably duration (1.5 hrs), the counts should likewise be roughly comparable across speakers.[10]

The second issue is that there are not enough tokens from each speaker to examine frequency effects at the speaker level, especially once phonological effects and word class have been taken into account. Moreover, it is difficult to disentangle the effects of word frequency and phonological context within a single speaker's data, as any given word will always have both a particular phonological context and a particular

[10] The use of corpus-internal measures of frequency has precedent in the literature, e.g Clark & Trousdale (2009). The speaker-internal approach adopted here is simply an extension of the corpus-internal approach, one which has the additional benefit of disentangling word frequency and phonological context.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Table 7. *Effects of frequency on F1 and F2 for each word class*

| | Effect of frequency (Hz/count) | p |
|---|---|---|
| (oh) F1 | −0.38 | 0.019 |
| (oh) F1 | (−0.03) | 1 |
| (o) F1 | 0.52 | 0.008 |
| (o) F2 | 1.72 | <0.001 |

interview frequency. These problems were addressed by pooling the data from all 17 speakers' conversational speech for the frequency analysis. This approach both increased the amount of data in the analysis and decoupled phonological context and frequency: given the speaker-specific frequency coding described above, words may vary in frequency coding across speakers. In order to pool the unnormalized formant data, it was necessary to correct for the gross formant magnitude differences across speakers; this was accomplished by including a Speaker random effect in the models described below.

Assuming older exemplar decay, usage-based models predict that high-frequency lexical items should be more advanced with respect to phonetic changes. In the present case, this means that high-frequency low back vowel words should be the first to show signs of shift towards New York-area realizations. Importantly, this change involves word class divergence: if shift occurs, (o) words should move fronter and lower in the vowel space, while (oh) words should shift higher and backer in this space. The effect of frequency is therefore expected to interact with word class. Higher-frequency (o) words are predicted to have higher F1 and F2 than low-frequency (o) items, while high-frequency (oh) words ought to have lower F1 and F2 than low-frequency items in that category.

To test each of these predictions, separate analyses of F1 and F2 were completed for each of (o) and (oh). Again, a model comparison procedure was used. For each analysis, the crucial comparison was between a model which includes only fixed phonological effects and a random effect of speaker, and a model which includes the same fixed phonological effects, the random effect of speaker, and a fixed effect of frequency. Results of the four analyses are summarized in table 7.

First consider (oh). Frequency has a negative effect on both formants, though this is only significant on the F1 dimension: higher frequency is associated with lower F1 values even after phonological factors have been taken into account. This means that high-frequency (oh) words are realized in a higher position in the vowel space than low-frequency words of this word class. For (o), however, frequency has a significant positive effect on both formants. Higher-frequency (o)s are associated with higher F1 and higher F2 values; that is, high-frequency (o) words are realized lower and fronter than their low-frequency counterparts. While these effects are quite small (possibly due to the unaccounted-for impact of the sort of cognitive factors that Pierrehumbert

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

discusses), they are consistent with the predictions of lexical and phonetic gradualness implied by the usage-based theory.

Finally, it is worth nothing that the frequency effects reported for conversational speech are consistent with the overall patterns of style shift shown by speakers across word list and minimal pair tokens. While higher-frequency items of both word classes are more advanced in the shift towards New York-area English in conversation, there is an asymmetry in the magnitude of these effects: high-frequency (o) items are more advanced with respect to frontness and height, while high-frequency (oh) items differ only in height; moreover, the effects are greater for (o), indicating that this word class has shifted to a greater extent. A similar pattern occurs in the read styles: for speakers who separate these vowels in the word list context, it is (o) which shows the greatest shift from minimal pair productions.

### *2.3.4 Judgment task*

The majority of speakers show evidence of having acquired a contrast between the (o) and (oh) word classes in spontaneous speech. Are these speakers aware of the distinction they have started to acquire? It seems intuitive that awareness of a feature would have some effect on its realization, though it is perhaps less clear which direction the influence will take. If a feature is stigmatized, then speakers might use less of it, but if the feature is not stigmatized, or if people see the feature as being associated with some identity that they view positively, then they might more quickly adopt it.

Typically, mergers are thought to be below the level of social awareness (Labov 1994: 324); that is, while speakers may be aware of how particular sounds undergoing a merger are realized, they are not consciously aware of mergers or distinctions as such. This view seems to be borne out by a lack of speaker comments about the low back vowels in the conversation portion of the interviews: when the discussion turned to linguistic features that differ between Canada and New York, no speaker offered up the pronunciation of *cot/caught*-type words as a feature that differs between their native and new dialect regions. One speaker mentioned the Brooklyn pronunciation of *dog*, producing this word with an extremely high vowel, but did not generalize this realization to other (oh) words, nor mention a difference between word classes. Of course, it may be that this feature is just not very salient compared to other dialect differences which might come up in such a conversation (e.g. Canadian Raising, or the discourse marker *eh*), or that lay speakers have a hard time articulating what this feature is.

The judgment task was carried out to more directly probe awareness of the low back contrast. In this task, speakers were asked to identify minimal pairs which they thought New Yorkers would produce differently and to imitate these productions where possible. The results of this task were not always conclusive. However, there are some speakers who clearly grasp that there is a (o)/(oh) distinction in the second dialect and some who are completely unaware of this difference.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Seven speakers display a strong awareness that there is a contrast between these two vowels in the ambient dialect, as well as an accurate, if exaggerated, grasp of the nature of the phonetic difference. GH, JC, JF, LW, NW and TM noted the difference for many of the (o)/(oh) pairs on the list, producing an extremely high back (and often lengthened) vowel for the (oh) word in each pair. LC is also aware of the contrast, but varies in where she locates the difference between Canadian English and New York English. For instance, she says that *caught/cot* are different in New York English, claiming that *caught* sounds like [kʊət], but for *don/dawn* and *odd/awed* she said that the difference is due to *don* and *odd* being produced, respectively, [dan] and [ad], with a very fronted low vowel. These seven speakers also made statements indicating awareness of a more general contrast beyond the individual differences between words on this list. These generalizations usually referenced orthography, e.g. LC's observation that 'a lot of the times just in general o's are a's, like dot com is [dat kam], like it's an *a* sound'.

LG, interestingly, seems to be aware of the difference, but for the most part gets the phonetics wrong. While she did pick out the low back vowel pairs as being produced differently by New Yorkers, she claimed that *talk* and *caught* are produced by locals as [tak] and [kat]. However, she does note that New Yorkers say *dog* like [dʊəg].

CW's responses are more difficult to interpret. The only pair she says would be different for New Yorkers is *caught/cot*, and she produces the right phonetic distinction, with *caught* having the higher backer realization. However, for the remainder of the pairs, she attempts both words with the exaggerated high back vowel, then the lower fronter vowel, before deciding that they are probably the same. A possible interpretation of this behavior is that while she is unaware that there is a general contrast, she does grasp that there is a wider range of acceptable pronunciations for this putatively single vowel category.

Four of the speakers pick out one or two words or word pairs as being different, but do not show awareness of a general contrast. BW, given the (distractor) pair *coal/call*, says people from New Jersey say [kwɑl], and points out a 'subtle elongation' of the vowel in *pawned* as compared with that in *pond*, but otherwise does not seem to generally grasp that there is a difference. PW says *caller* is more 'drawn out' than *collar*, but produces the first word with a much more fronted vowel. SS says *tall* may be different from *doll*, but doesn't point out any other pairs. VJ says *doll* may be produced with a fronter, more 'drawn out' vowel, but otherwise does not spot any low back differences. Finally, BW, DB, ES and EW betray no awareness of a difference in the low back vowels, either phonological or phonetic. These speakers completely glossed over the low back vowel pairs in doing this task (and thus did not produce imitation tokens of these), focusing instead on features such as *r*-lessness in words like *higher/hire.*

In summary, four speakers seem to be clearly unaware of the low back vowel contrast, seven speakers appear to have an accurate grasp of the general contrast as well as its phonetic realization, and the remaining six speakers fall somewhere in between. This variation in awareness of the feature across speakers does not, however, relate in any obvious way to the variation across speakers in realization of the contrast in spontaneous

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Table 8. *Awareness of the low back vowel contrast vs. realization of that contrast*

| Awareness of contrast | (o)/(oh) same | (o)/(oh) different |
|---|---|---|
| Unaware | ES | BW, DB, EW |
| Maybe aware | CW, PW, VJ | BK, LG, SS |
| Aware | NW, TM | GH, JC, JF, LC, LW |

speech, as shown in table 8; it so happens that five of seven aware speakers realize the contrast, but so do three of the four unaware speakers.

## 3 Discussion

The study of low back vowel realization among mobile Canadians reported here demonstrates that new contrasts may be acquired by speakers later in life. It must be noted, however, that these speakers show remarkable stability in their low back vowel systems. This is most clearly evident in the minimal pair results: nearly all speakers are merged in production and perception in this context. Where speakers do make a significant distinction between (oh) and (o) in spontaneous or word list speech, the phonetic difference is quite subtle compared with the robust distinction made in New York-area English.

That said, the majority of speakers do show evidence of having acquired a distinction between (o) and (oh) in their spontaneous speech on at least one phonetic dimension. That is, these speakers show phonetic variation in these vowels that cannot be attributed to phonological context alone, but can be at least partially explained by word class membership in the ambient dialect. This change, where it has occurred, seems to be phonetically and lexically gradual: there remains extensive overlap between the two word classes, with higher word frequency being associated with more New York-like phonetic realizations.

As noted in section 1, these results can be brought to bear on the issue of phonological representation: which kind of model best accounts for how these speakers have changed their vowel production? An abstractionist account of these results might be that these speakers have managed to change their underlying forms for some relevant lexical items to reflect the contrast in their new dialect. Words such as *cot* and *caught,* previously represented identically as /kɑt/ and /kɑt/, are now stored as /kɑt/ and /kɔt/, respectively. The realization of each of these new categories – in particular, the magnitude of the phonetic distance between them – is not clearly predicted; all we know is that they ought to be different. The subtlety of the surface distinctions evident in the data is more easily accommodated in usage-based theory, and indeed predicted: contrast is not achieved in a featural quantum leap, but gradually, via the addition of exemplars at the word level, which ultimately lead to a more general divergence at the word class level.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Further support for a usage-based account comes from the frequency effects observed in this data. High-frequency (oh) words are higher than other (oh) words, while high-frequency (o) words are lower and fronter, indicating that high-frequency items are in the vanguard of divergent shift within their respective word classes in the low back vowel spaces of these speakers. These facts indicate a lexically gradual shift towards the new variety: speakers hear high-frequency words more often, meaning that they acquire new dialect exemplars of these words at a faster rate, which results in the representations (and thus productions) of these words shifting before those of less frequent words. These results are difficult to accommodate within the abstractionist account; the best it can do is posit lexical exceptions which generate these results, but in such an account the fact that these exceptions are structured in terms of frequency would be mere coincidence.

Moreover, the lack of a relationship between awareness of the ambient contrast and production of this contrast in spontaneous speech as revealed by the judgment task is difficult to account for within an abstractionist model. Speakers who produce a distinction but are unaware of the distinction are a particular problem for this view: such speakers would seem to have acquired a 'covert contrast' that is for some reason not accessible to intuition, even though it is formally indistinguishable from any other feature-based contrast in the system. The dissociation of production and intuition is less problematic in usage-based theories, where new productions are based on clouds of remembered tokens, whether or not new abstract category labels are present.

While most of the speakers make a significant distinction between (o) and (oh) in spontaneous speech, none of these speakers exhibit that distinction in minimal pair speech. This is, on the face of it, strange behavior for a minimal pair task. Minimal pairs highlight possible contrasts, and are thus the context in which contrasts – even marginal ones – are most likely to surface. In Labov's (1966) study of (r) on the Lower East Side, for example, speakers contrasted word pairs like *sauce/source* most consistently in the minimal pair context, using more coda (r) in this style versus the connected speech styles. Even in cases of near-merger, where speakers do not themselves perceive the difference in their speech, the marginal contrast will reveal itself in the production part of minimal pair tests (Labov *et al.* 1991). The Canadians in this study, however, behave in the opposite way: the marginal distinction in their conversational speech is eradicated in just the context in which it should be most likely to appear.

An explanation for this patterning may come from considering just what minimal pair tasks are meant to elicit. Labov (1966: 152) sets minimal pair tasks (along with word lists) apart from the connected speech styles he analyzes, noting that the citation styles are better taken as an indication of 'phonic intention, illustrating the norms of the speaker, in part, rather than a reliable indication of performance'. In the case of the New Yorkers Labov interviewed, the norm which was illustrated in minimal pair speech was (r)-fulness; this reflected the local change in progress towards the wider norm of realizing coda (r). Labov's speakers may not have consistently produced (r) in their connected speech, but at some level they knew that they should do so.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

The expatriate Canadians in this study find themselves in a very different social context. They are not natives of a speech community undergoing change, but newcomers to a community with stable, though different, norms. However, these new norms do not seem to be adopted as such by the mobile speakers, even though their conversational speech shows evidence of their influence. Instead, it seems that the Canadian speakers maintain their first dialect norms for low back vowel realization. These findings have important methodological implications for the study of merger and split, especially among speakers in dialect contact situations: the sociolinguist cannot safely rely on the minimal pair test as the style which will bring out contrast; more extensive analysis of conversational data may be necessary to reveal a subtle distinction.

*Author's address:*
*Department of Linguistics*
*Georgetown University*
*1437 37th St NW*
*Washington, DC 20057*
*USA*
*jn621@georgetown.edu*

## References

Archangeli, Diana. 1988. Apects of underspecification theory. *Phonology* 5, 183–207.
Baayen, R. Harald. 2008. *Analyzing linguistic data: A practical introduction to statistics*. Cambridge: Cambridge University Press.
Baayen, R. Harald, Richard Piepenbrock & Hedderik van Rijn. 1993. *The CELEX lexical database*. Linguistic Data Consortium, University of Pennsylvania.
Bates, Douglas & Deepayan Sarkar. 2008. *lme4: Linear mixed-effects models using s4 classes*. http://cran.r-pro ject.org.
Bloomfield, Leonard. 1926. A set of postulates for the science of language. *Language* 2(3), 153–64.
Boberg, Charles. 2008. English in Canada: Phonology. In Edgar W. Schneider (ed.), *Varieties of English: The Americas and the Caribbean*, vol. 2, 144–60. Berlin: Mouton de Gruyter.
Bybee, Joan. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.
Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. New York: Harper & Row.
Church, Barbara A. & Daniel L. Schacter. 1994. Perceptual specificity of auditory priming: Implicit memory for voice intonation and fundamental frequency. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20, 521–33.
Clark, Lynn & Graeme Trousdale. 2009. The role of frequency in phonological change: Evidence from TH-fronting in east-central Scotland. *English Language and Linguistics* 13(1), 33–55.
Clements, George N. 1985. The geometry of phonological features. *Phonology Yearbook* 2, 225–52.
Clements, George N. & Elizabeth Hume. 1995. The internal organization of speech sounds. In John A. Goldsmith (ed.), *The handbook of phonological theory*, 245–306. Cambridge, MA: Blackwell.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Cole, Ronald A., Max Coltheart & Fran Allard. 1974. Memory of a speaker's voice: Reaction time to same- or different-voiced letters. *Quarterly Journal of Experimental Psychology* 26, 1–7.
Gafos, Adamantios. 2002. A grammar of gestural coordination. *Natural Language and Linguistic Theory* 20, 269–33.
Goldinger, Stephen D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251–79.
Goldsmith, John A. 1979. The aims of autosegmental phonology. In Daniel A. Dinnsen (ed.), *Current approaches to phonological theory*, 202–22. Bloomington: Indiana University Press.
Herzog, Marvin. 1965. *The Yiddish language in northern Poland*. Bloomington and The Hague: Mouton & Co.
Hintzman, Douglas L. 1986. 'Schema abstraction' in a multiple-trace memory model. *Psychological Review* 93, 411–28.
Hintzman, Douglas L., Richard A. Block & Norman R. Inskeep. 1972. Memory for mode of input. *Journal of Verbal Learning and Verbal Behavior* 11, 741–9.
Jakobson, Roman. 1962. *Selected writings*, vol. 1. The Hague: Mouton & Co.
Johnson, Daniel Ezra. 2010. *Stability and change along a dialect boundary: The low vowels of southeastern New England*. Publications of the American Dialect Society 95. Durham, NC: Duke University Press.
Johnson, Keith. 1997. Speech perception without speaker normalization. In Keith Johnson & John W. Mullennix (eds.), *Talker variability in speech processing*, 145–66. San Diego, CA: Academic Press.
Kager, René. 1999. *Optimality Theory.* Cambridge: Cambridge University Press.
Kučera, Henry & W. Nelson Francis. 1967. *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
Labov, William. 1966. *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics, 1st edition.
Labov, William. 1994. *Principles of linguistic change: Internal factors*. Cambridge, MA: Blackwell.
Labov, William. 2010. *Principles of linguistic change: Cognitive and cultural factors*. Cambridge, MA: Blackwell.
Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology, and sound change: A multimedia reference tool*. Berlin: Mouton de Gruyter.
Labov, William, Mark Karen & Corey Miller. 1991. Near-mergers and the suspension of phonemic contrast. *Language Variation and Change* 3, 33–74.
Labov, William, Malcah Yaeger & Richard Steiner. 1972. *A quantitative study of sound change in progress*. Philadelphia, PA: US Regional Survey.
Ladefoged, Peter. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Cambridge, MA: Blackwell.
Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical perspectives*. Stanford, CA: Stanford University Press.
Langacker, Ronald. 2000. A dynamic usage-based model. In Michael Barlow & Susanne Kemmer (eds.), *Usage-based models of language*, 1–63. Stanford, CA: CSLI Publications.
Malkiel, Yakov. 1967. Every word has its own history. *Glossa* 1, 137–49.
Mullennix, John W., David B. Pisoni & Christopher S. Martin. 1988. Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America* 85, 365–78.
Norris, Dennis, James McQueen & Anne Cutler. 2003. Perceptual learning in speech. *Cognitive Psychology* 47, 204–38.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press

Nygaard, Lynne C. & David B. Pisoni. 1998. Talker-specific learning in speech perception. *Perception and Psychophysics* 60, 355–76.
Paul, Hermann. 1880. *Prinzipien der Sprachgeschichte*. Halle: Niemeyer. [English translation of 2nd (1886) edition: *Principles of the history of language*, trans. H. A. Strong. College Park: McGrath Publishing Company, 1970.]
Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–42.
Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition, and contrast. In Joan Bybee & Paul J. Hopper (eds.), *Frequency and the emergence of linguistic structure*, 137–57. Amsterdam: John Benjamins.
Pierrehumbert, Janet. 2002. Word-specific phonetics. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory phonology 7*, 101–39. Berlin: Mouton de Gruyter.
Pierrehumbert, Janet. 2003. Probabilistic phonology: Discrimination and robustness. In Rens Bod, Jennifer Hay & Stefanie Jannedy (eds.), *Probabilistic linguistics*, 177–228. Cambridge, MA: MIT Press.
Pierrehumbert, Janet. 2006. The next toolkit. *Journal of Phonetics* 34, 516–30.
Pinheiro, Jose C. & Douglas M. Bates. 2000. *Mixed-effect models in S and S-Plus*. New York: Springer.
Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.
Schacter, Daniel L. & Barbara A. Church. 1992. Auditory priming: Implicit and explicit memory for words and voices. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18, 915–30.
Steriade, Donca. 1995. Underspecification and markedness. In John A. Goldsmith (ed.), *The handbook of phonological theory*. Cambridge, MA: Blackwell: 114–75.
Trubetzkoy, Nikolai S. 1969. *Principles of phonology*, trans. Christiane A. M. Baltaxe. Berkeley and Los Angeles: University of California Press. Originally published 1939 as *Grundzüge der Phonologie*. Göttingen: Vandenhoeck & Ruprecht.
Trudgill, Peter & Nina Foxcroft. 1978. On the sociolinguistics of vocalic mergers: Transfer and approximation in East Anglia. In Peter Trudgill (ed.), *Sociolinguistic patterns in British English*, 69–79. London: Edwin Arnold.
Vaux, Bert & Justin Cooper. 1999. *Introduction to linguistic field methods*. Munich: Lincom Europa.
Wedel, Andrew. 2004. Category competition drives contrast maintenance within an exemplar-based production/perception loop. In John A. Goldsmith & Richard Wicentowski (eds.), *Proceedings of the seventh meeting of the ACL Special Interest Group in Computational Phonology*, vol. 7, 1–10. ACL.
Wedel, Andrew. 2006. Exemplar models, evolution and language change. *The Linguistic Review* 23, 247–74.

https://doi.org/10.1017/S1360674313000051 Published online by Cambridge University Press