The Spread of Raising: Opacity, Lexicalization, and Diffusion

University of Pennsylvania Working Papers in Linguistics

Volume 14
Issue 2 *Selected Papers from NWAV 36*

Article 11

11-22-2008

The Spread of Raising: Opacity, Lexicalization, and Diffusion

Josef Fruehwald
University of Pennsylvania, joseff@alumni.upenn.edu

The Spread of Raising: Opacity, Lexicalization, and Diffusion

Josef Fruehwald

1 Introduction

The phenomenon commonly referred to as Canadian Raising involves the raising of the nuclei of /ay/ from a low to mid position before voiceless consonants. It was first reported on by Joos in 1942 as a change taking place in heartland Canadian English. Since that time, raising has been found throughout many areas in the Northern United States (Labov, Ash, and Boberg, 2006). In most dialects where raising is present, it stands in a counter-bleeding relationship with intervocalic flapping of unstressed /t/. In these dialects, the distinction between *writer* and *rider* is maintained in the quality of the vowel rather than in the voicing of the stop. As such, raising is frequently cited as a typical example of phonological opacity.

This paper explores the apparent lexicalization of raising, and following diffusion /ʌy/ to new contexts, specifically in Philadelphia. A number of papers on raising in Canada and the North Central and Inland North of the United States have identified lexical exceptions where low nuclei occur when raised nuclei are expected (Chambers, 1973, 1989; Vance, 1987). This research specifically investigates the distribution of raised nuclei before voiced consonants. The distribution of raised nuclei before voiced segments is not entirely erratic, but is not phonologically regular either. Rather, it appears as if the raised vowel /ʌy/ has lexicalized due to raising's opaque applications and is spreading to similar environments.

2 Raising's Phonology

The phonological description of raising as described in the literature, both its conditioning factors and its location in the phonological system, are slightly more complicated than simply "raise before voiceless." I will review them here briefly.

2.1 Conditioning

Chambers (1973, 1989) in Canada and Vance (1987) in upstate New York have found that following primary, and occasionally secondary stress will block raising ([tʌyˈ.tən] versus [tay.tæˈ.nɪk]) Given these facts, raising must interact with syllabicity. Chambers (1973) explained the data by saying that the following voiceless segment must be ambisyllabic with /ay/ to trigger raising, and Jensen (2000) explained it by saying that the voiceless segment must be in the same foot as /ay/ to trigger raising.

Raising also exists at a fairly abstract level in the phonology, as it frequently underapplies at morpheme and word boundaries. As such, the presence or absence of raising can frequently be used to determine a speaker's morphological analysis of words. For example, raising in *high school*, *high-chair* and *bicycle* is a good indication that these words have been reanalyzed as monomorphemic. Using Canadian raising as an example case for Stratal OT, Bermudez-Otero (2003) postulated that raising occurred at the stem level given the *eyeful*∼*Eiffel* distinction.

2.2 Opacity

As mentioned above, raising exists in a counter-bleeding relationship with flap formation. When an unstressed /t/ is voiced to a [ɾ], raising continues to apply opaquely. This relationship between raising and flap formation was first pointed out by Joos (1942). He also observed that there were two groups of school children, one which applied raising opaquely, and one which didn't. For the second group, a distinction would be maintained between *write* and *ride* ([rʌyt]∼[rayd]), but not between writer and rider ([rayɾɝ]∼[rayɾɝ]). Today, this second group of speakers no longer exists, and all speakers have the opaque pattern (Chambers, 1973).

2.3 Acquisition

The grammatical conditioning of raising poses a particular problem for learners. There is both grammatically conditioned underapplication by stress shifting affixes, and grammatically conditioned overapplication via flap formation, leading to the appearance of a contrast between [ay]∼[ʌy]. In order to arrive at the conclusion that the contrast between the words *writer* and *rider*, for instance, is crucially on the underlying voicing of the flap and not on the vowel, a child must carry out a morphological analysis of the two words to arrive at /rayt + r/ and /rayd + r/. The situation is further complicated for a child in the case of monomorphemic words where raising also applies opaquely, such as *miter*. Furthermore, the phonetic correlate of voicelessness (longer VOT) is greatest in stressed syllable onsets, exactly the following context which ought to block raising.

As discussed in Hayes (2004), children acquire most of their phonological knowledge earlier than they do their morphological knowledge. Hayes' proposal is that children at a young age may have a phonemic contrast between two grammatically conditioned allophones, but once they begin to perform the proper morphological analysis, the distribution of the two phones will become entirely predictable, and thus allophonic.

2.4 Lexicalization

There have been a handful of papers suggesting that raising has been lexicalized. This was, in fact, the initial prediction of Joos (1942). He suggested that since /ay/ and (in Canada) /aw/ behaved differently before voiceless consonants (centralizing their nuclei), than any other vowel (shortening their duration), this opened the door for phonemicization. Mielke et al. (2003) argued from a theoretical perspective that the predictable distribution of [ay]∼[ʌy] is a historical relic of transparent allophony rather than a synchronic opaque allophony. This is an interesting point to make, since if [ay]∼[ʌy] were to phonemicize, it would, in all likelihood, happen silently until their distributions began to change as a result of other language changes.

Vance (1987) has the first detailed report of observed lexicalization of raising. He gathered data from the North Central and Inland North from himself, his mother and a high school friend, asking them for their judgments as to whether a word had a raised or low nucleus. For most words, all three speakers agreed, and they were in alignment with previous descriptions of raising. They diverge greatly from previous accounts of raising, however, by raising /ay/ before /r/. They also had raising before a few voiced stops in *spider*, *cider*, *idle*, and *tiger*. Importantly, it was not all tokens before /r/, or a flapped /d/ or an unstressed stop which experienced raising. In fact, one of Vance's informants even had a distinction between *idol* and *idle*, leaving the distribution of [ʌy] phonologically unpredictable. In Ann Arbor, also in the Inland North, Dailey-O'Cain (1997) also found significant raising before /r/, as well as before sequences consisting of a nasal followed by a voiceless consonant.

A more recent study done by Hall (2005) in Meaford, Ontario found a number of lexical exceptions while performing careful phonetic analysis. By accounting for the preceding segment, she could explain most exceptions. However, rather than the phonological or phonetic properties of the initial segment effecting raising, she attributed it to a lexical neighborhood effect. That is, words which were more similar to each other were more likely to have the same [ay]∼[ʌy] variant.

3 Current Study

The goal of this study was to investigate the extent to which raising has spread to new pre-flap contexts and other pre-voiced contexts as well as to see if the particular conditioning environments described in section 2.1 remain productive.

3.1 Raising in Philadelphia

Raising was first observed in Philadelphia in the 1970s during Bill Labov's LCV studies and was identified as an incipient male-led change (Labov, 2001). Since that time, it has become a well-

studied sociolinguistic variable referred to in the literature as (ay0). Conn (2005) found that women were also partaking in raising in pre-voiceless contexts, but men now led backing of (ay0), hence the title of his dissertation: *Of Moice and Men*. In subjective reaction tests, he found that women were downgraded for the backness of their (ay0), and men were upgraded. In studying high school girls in South Philadelphia, Wagner (2007) found that girls more invested in a "hair and make-up" construct of their femininity had less backed (ay0) than girls who tried to be "one of the guys" or who played competitive sports.

3.2 Methods

For my fieldwork, I conducted interviews at an all-male high school in Philadelphia, which I will refer to by the pseudonym Archbishop Cahill High School. My data collection consisted of a 15 to 20 minute interview with each speaker, which included 14 semantic differential questions, a recitation of Little Miss Muffet, two minimal pair tests and a word list. After data collection, I measured every token of [ay] and [ʌy] at F1 maximum for 12 speakers, aged 14 through 18. Each token was coded for environmental data, including place, manner and voicing of the following segment, the place and nasality of the preceding segment, the structure of the coda and how many following syllables there were, and the stress of the following syllable. After checking for errors, I also coded each token for the presence of following morpheme boundaries. In total, 1751 tokens were measured and coded.

3.3 Results

3.3.1 Speaker Data

Figure 1 displays the vowel distributions of three speakers from this study, identified by their pseudo-initials KT, DJB and JP. Tokens which would be phonologically predicted to be raised are black triangles, and those predicted to be low are in grey circles. As can be seen in KT's distribution, there are a number of tokens predicted to be raised amongst the predominantly low cluster, and even more tokens predicted to be low amongst the raised cluster. While there does appear to be a raised concentration and a low concentration in KT's distribution, the over all pattern looks fairly continuous, similar to, but not as continuous as, the distributions in Hall (2005). This may immediately call into question the claim that there are two phonemes, or quasi-phonemes, in these speakers' inventories. However, speaker DJB has a much clearer raised cluster, and speaker JP has a rather extraordinary distribution with two very clearly defined raised and low clusters without much phonetic overlap. Even for JP, some tokens which would be expected to be low are raised, and vice versa.

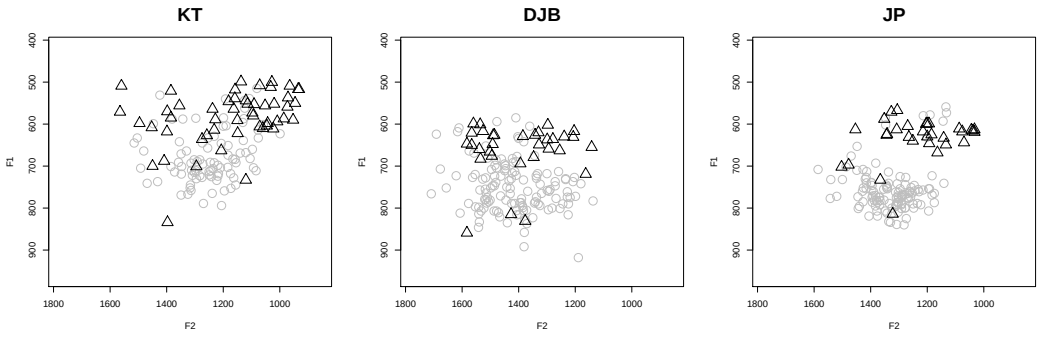

Figure 1: Vowel plots in Hz, with words expected to be raised in black

3.3.2 Regression Analysis

I performed a linear regression analysis on the data to find the significant factors in determining vowel height. The dependent variable in the regression was F1 as a measure of vowel height. To

control for variation between speakers, each speaker was entered as a factor group in the analysis first.

The first analysis performed looked only at the phonological environment. The factors that were included were: Following Voice, Manner and Place of Articulation, Following Primary Stress, Morphemic Structure, Syllabic Structure and Onset Place.

When entering all of these factors into the regression, the only one to produce a significant result was a following voiceless consonant, which had a coefficient of -94.6 on a constant of 722.8 at a significance of $p = 0.01$. The negative value of the coefficient indicates an effect of decreasing F1, raising the vowel. As non-significant factors were removed, a few more factors became significant, as displayed in Figure 2. The y-axis is aligned to indicate vowel height. Factors above the constant (dot-dash) line had an effect of raising the vowel, and factors below had an effect of lowering the vowel. The effect of a following voiceless segment is displayed as the dashed line, for means of comparison to other factors.

A following voiceless segment has the expected effect of raising the vowel, and has the largest effect of all the significant factors. The raising effect that a following apical segment and a labial onset have may simply be fine grained phonetic effects. Unsurprisingly, a following nasal and /r/ have an effect of having lower vowels, as does a following morpheme boundary to a lesser degree.

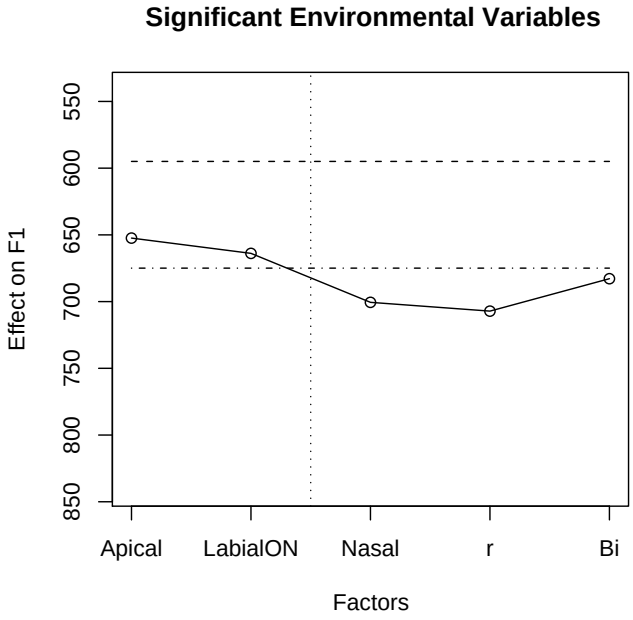

Figure 2: Factor effects on F1 for phonological variables

Next, I entered factor groups that controlled for the particular onset of the following syllable, that is, the particular syllable sequence /ay.p/, /ay.b/ etc. Even though /ay.t/ and /ay.d/ would both appear as flaps on the surface, they were given separate factor groups in the regression. Importantly, the abstract phonological properties of these syllable sequences are already controlled for in the regression with the factor groups for voice, place and manner of the following segment, syllable structure and stress. Significant effects of these syllable sequence factors indicate effects of particular segments, and already call into question the phonological regularity of vowel height. The results of that regression analysis are presented in Figure 3. The effect of a following apical segment has dropped out of the analysis, replaced by following underlyingly voiced, and underlyingly voiceless flaps. Interestingly, the effect of an underlying voiced flap is *greater* than an underlyingly voiceless flap.

Next, a linear regression was carried out for each word which was represented by multiple

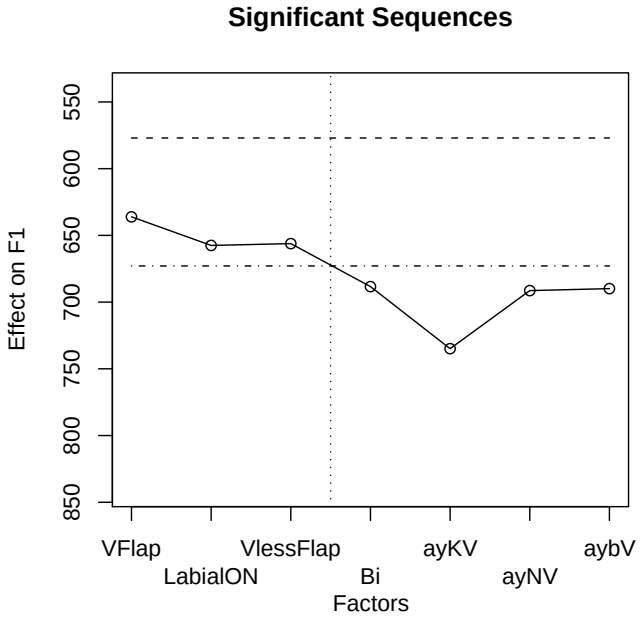

Figure 3: Factor effects on F1 for syllable sequences

tokens in the corpus. Included in the regression was a factor for the particular word, factors for the environmental variables listed above as well as the syllable sequence factors. Even with all of the phonological factors, and even some lexical factors controlled for, a number of words returned significant effects on vowel height. They are presented in Figure 4, along with the constant and effect of a following voiceless segment from their particular regression. The significant effects of *spider* and *cider* appear to agree with the findings in Vance (1987), which also found these words to have raised nuclei.

3.3.3 Individual Analysis

While the regression analysis demonstrated the lexical effects of various words, the way individual speakers behave in regards to these effects is of some interest. Whether or not speakers produce these words categorically raised, categorically low, or vary between the two, or whether nor not there is great variation between speakers remain open questions. To approach answers to these questions, an adequate method of describing individual speaker behavior needs to be used. Impressionistic coding of every token as raised or low is one possible course of action. However, the confusability of tokens on the near edges of each distribution combined with interference from my own native phonetic processing, and phonological and lexical expectations could confuse the output.

Instead, a Perl script was used to carry out k-means clustering on each speaker's vowel distribution. This clustering algorithm requires a researcher to specify a number of clusters, and iteratively improves the clustering, under the assumption that errors are normally distributed. It took the means of each cluster as were coded by expected class given the phonological environment, then re-classified each token based on its Euclidean distance from each mean. Then, the means were re-calculated based on this re-clustering, and the process continued until means were reached for which no tokens would be reclassified in the next iteration. Then, for each cluster, tokens which were more than one standard deviation from the mean in the direction of the other cluster were trimmed, leaving two clusters which contained tokens I could say were produced raised or low with a good deal of confidence. Figure 5 shows the vowel plots for one speaker's data as they were processed in this manner. After processing all 12 speaker's data in this way, one speaker's data was eliminated from

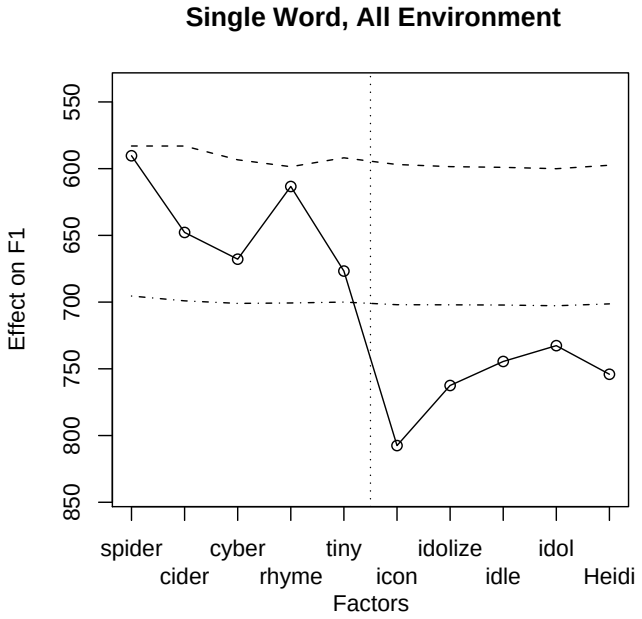

Figure 4: Lexical items' effects on F1

further analysis because his particular vowel distribution caused the algorithm to behave erratically.

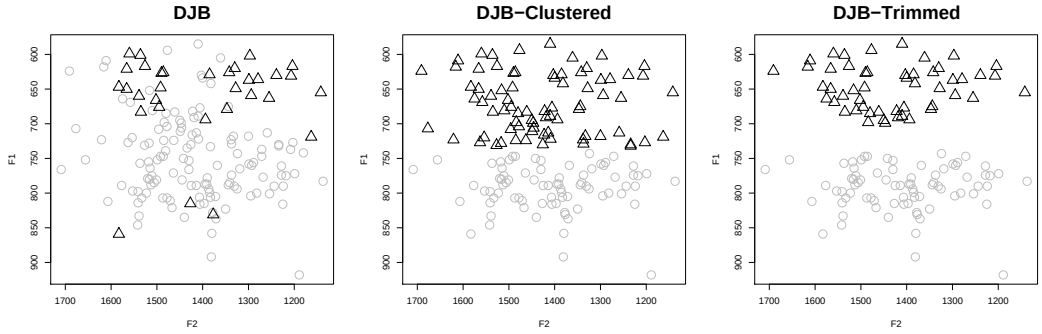

Figure 5: Clustering process for speaker DJB

Now that all tokens for each speaker were classified into raised and low classes, words which were produced unexpectedly raised were identified. To account for the possibility of disfluency, measurement error or clustering error, a word was only considered raised if at least 4 of the 11 speakers produced it raised. The results of this cluster analysis in Table 1 agrees with the results of the regression analyses. *Spider* and *cider* are produced raised unanimously by all 11 speakers, but *cyber* and *tiny* are only produced raised by 6 of 11 speakers, reflecting their smaller coefficients in the regressions. *Rhyme* is not produced raised by 4 or more speakers, and in the regression analyses, it had a much larger standard error than other lexical items with similar coefficients. It may be that one particular outlier in the corpus gave *rhyme* a significant coefficient in the regression that is not reflective of the communal behavior.

In order to illustrate internal speaker variation and variation between speakers, Tables 2, 3, and 4 display the number of speakers who produced each word categorically raised, categorically low, and the distributions for those who produced them mixed. For *spider* the unanimity that it is produced

raised is also reflected in categorically raised productions for 10 out of 11 speakers. *Tiny* and *cyber* reflect more intermediate patterns, with some speakers producing them categorically raised, others producing them categorically low, and many producing both.

Word	Speakers
spider	11/11
cider	11/11
beside	6/11
cyber	6/11
tiny	6/11
idol	5/11
tidal	5/11
dinosaur	4/11
Heinz	4/11
nine	4/11
Snyder	4/11
tidy	4/11

Table 1: Words and number of speakers who produced them raised

Spider

Behavior	Speakers	Speaker	Raised	Low
Categorical Raised	10	RJM	3	2
Categorical Low	0			
Mixed	1			

Table 2: Speaker behavior: *spider*

Cyber

Behavior	Speakers	Speaker	Raised	Low
Categorical Raised	2	SEP	1	3
Categorical Low	4	JAF	3	2
Mixed	5	SPM	1	2
		FJC	1	1
		RJM	1	1

Table 3: Speaker behavior: *cyber*

3.3.4 Conclusions

Clearly, the phonological conditioning described in section 2.1 cannot adequately account for the pattern off raising observed in the data from these 12 speakers from Philadelphia. The presence of a lexical effect upon raising, even after controlling for the phonological environment seems to indicate that raising has phonemicized. The persistent effect of a following voiceless consonant could be the result of all words with /ay/ and a following voiceless consonant being specified with /ʌy/. However, it need not be the case that an available phoneme /ʌy/ excludes the possibility of a productive process in the phonology that raises /ay/ nuclei before voiceless consonants. For this reason, low frequency, novel, or nonce words may still be regularly produced with [ʌy], even though they were introduced after lexical specification took place.

Tiny

Behavior	Speakers	Speaker	Raised	Low
Categorical Raised	1	SEP	1	4
Categorical Low	5	JGB	2	2
Mixed	5	SPM	3	2
		BGS	1	1
		KT	4	1

Table 4: Speaker behavior: *tiny*

3.4 Lexicalization and Diffusion

How these lexical specifications of raised nuclei should be analyzed depends upon the history of raising. One proposed history has been that raising is a residue of the Great Vowel Shift lowering of Middle English /ʌy/ to Modern English /ay/ (Gregg, 1973; Labov, 1963, 1965). Another suggested history is that raising is a phonetically felicitous development, thus allowing for its independent innovations in different dialects (Chambers, 1973; Joos, 1942; Moreton, 2004; Moreton and Thomas, 2004; Thomas, 1991, 2000).

If raising in Philadelphia has its origins in the Great Vowel Shift and subsequent dialect contact, then it is not inconceivable that the particular lexical effects found here came into the dialect at the same time as raising. In this case, a contemporary account of diffusion would not be necessary, as *spider* would simply never have experienced lowering from /ʌy/ to /ay/. This account may be the simplest way to explain how *spider* and *cider* would both be selected to have /ʌy/ in two different dialect areas: Philadelphia (this study), the Inland North (Dailey-O'Cain, 1997; Vance, 1987). However, this doesn't seem to be plausible for Philadelphia. The raising of /ay/ before voiceless consonants was a new and vigorous change in Philadelphia when first observed in the 1970s (Labov, 2001).

If these lexical exceptions did not come into the dialect as-is, then they must have developed from a previously regular [ay]∼[ʌy] allophony. Looking at the words which have already experienced diffusion, as indicated by the regression analyses and the individual analysis, an explanation suggests itself. The words with the largest coefficients and most agreement have /ʌy/ followed by a surface flap. Most of the rest have a similar form of an open syllable before an unstressed stop. If the lexical items selected for diffusion are indicators of the motivation for diffusion, then it would seem it is raising's opaque application as discussed in section 2.2 and the challenges involved in acquiring that opacity, as discussed in section 2.3 that have led to the diffusion of /ʌy/ to new contexts. Importantly, I believe that it is because /ʌy/ has its historical roots in a predictable allophony that a process of phonetically abrupt diffusion of /ay/ to /ʌy/ is observed here, rather than regular, phonetically gradual change from [ay] to [ʌy].

An interesting observation to make about this case of diffusion is that the new contexts in which /ʌy/ appear lack the initial phonetic conditioning that gave rise to the predictable allophony in the first place. Rosenfelder (2007) investigated the phonetic nature of raising in Victoria, B.C., particularly taking into account the *Offglide Peripheralization* hypothesis of Moreton and Thomas (2004). She found that while the nuclei of [ʌy] before flapped /t/ were raised more than before other, surface voiceless segments, the offglides before flapped /t/ were more similar to other surface voiced segments. Before surface voiceless segments, offglides were still peripheralized. It seems then, that it is not the phonetic felicity for /ʌy/ in a word that selects it for diffusion, but rather its phonetic "shape." The lexical neighborhood effects found in Hall (2005) further suggest that this is the way in which this diffusion is progressing.

4 Discussion

The results found here don't necessarily speak to the status of raising in any other region where raising takes place. If raising can arise in many dialects due to phonetic pressures exerted by a

following voiceless segment, then there is no need to treat it as a monolithic phenomenon. The data across studies even indicates that the status of raising in any given dialect can vary in any number of ways. Dailey-O'Cain (1997) and Vance (1987) both reported raising before /r/ in the North Central and Inland North, while Hall (2005), Rosenfelder (2007) found none in Canada and the current study found none in Philadelphia. Dialects vary in their phonetic character as well. The data in Hall (2005) form a fairly continuous distribution, while the data in Rosenfelder (2007) and this study form more clearly defined raised and low classes. However, both Hall (2005) and Rosenfelder (2007) show that raised vowels tend to be front of [ay], whereas many speakers in this study, and others in Philadelphia (Conn, 2005; Wagner, 2007) had their raised vowels directly above, or back of [ay].

Given this variability between studies, phonetic or phonological descriptions of raising ought to be defined by a particular dialect, and then supported with careful observations of that dialect.

References

Bermudez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In *Variation within Optimality Theory: proceedings of the Stockholm Workshop in 'variation within Optimality Theory'*, ed. Jennifer Spenader, Anders Eriksson, and Osten Dahl, 25–36.

Chambers, J. K. 1973. Canadian Raising. *Canadian Journal of Linguistics* 18:131–135.

Chambers, J. K. 1989. Canadian Raising: blocking, fronting, etc. *American Speech* 64:75–88.

Conn, Jeffery. 2005. Of 'moice' and men: The evolution of a male-led sound change. Doctoral Dissertation, University of Pennsylvania.

Dailey-O'Cain, Jennifer. 1997. Canadian raising in a midwestern U.S. city. *Language Variation and Change* 9:107–120.

Gregg, Robert J. 1973. The diphthongs əi and ai in Scottish, Scotch-Irish and Canadian English. *Canadian Journal of Linguistics* 18.

Hall, Kathleen Currie. 2005. Defining Phonological Rules over Lexical Neighbourhoods: Evidence from Canadian Raising. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, ed. et al. John Alderete, 191–199.

Hayes, Bruce. 2004. Phonological Acquisition in Optimality Theory: The Early Stages. In *Constraints in Phonological Acquisition*. Cambridge University Press.

Jensen, John T. 2000. Against Ambisyllabicity. *Phonology* 17:187–235.

Joos, Martin. 1942. A phonological dilemma in Canadian English. *Language* 18:141–144.

Labov, William. 1963. The social motivation of a sound change. *Word* 19.

Labov, William. 1965. On the mechanism of linguistic change. *Georgetown Monographs on Language and Linguistics* 18:91–114.

Labov, William. 2001. *Principles of linguistic change. Volume 2: Social Factors. Language in Society*. Oxford: Blackwell.

Labov, William, Sherry Ash, and Charles Boberg. 2006. *The Atlas of North American English. Phonetics, Phonology and Sound Change*. Berlin: Mouton de Gruyter.

Mielke, Jeff, Mike Armstrong, and Elizabeth Hume. 2003. Looking through opacity. *Theoretical Linguistics* 29:123–139.

Moreton, Eliott, and Erik R. Thomas. 2004. Origins of Canadian Raising in voiceless-coda effects: a case study in phonologization. In *Papers in Laboratory Phonology 9*, ed. J. Cole and J. I. Hualde. Cambrige University Press.

Moreton, Elliott. 2004. Realization of the English postvocalic [± voice] contrast in F1 and F2. *Journal of Phonetics* 32.

Rosenfelder, Ingrid. 2007. Canadian Raising in Victoria, B.C.: An Acoustic Analysis. *Arbeiten aus Anglistik und Amerikanistik* 32(2):257–284.

Thomas, Erik. 1991. The Origin of Canadian Raising in Ontario. *Canadian Journal of Linguistics* 36:147–170.

Thomas, Erik. 2000. Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. *Journal of Phonetics* 28:1–25.

Vance, Timothy. 1987. 'Canadian Raising' in some dialects of the northern United States. *American Speech* 62:195–210.

Wagner, Suzanne E. 2007. We act like girls and we don't act like men: The use of the male-associated variable (ay0) in South Philadelphia. In *Penn Working Papers in Linguistics*, ed. Tatjana Scheffler, Joshua Tauberer, Aviad Eilam, and Laia Mayol, volume 13.1, 393–406. Philadelphia: University of Pennsylvania Press.

Department of Linguistics
University of Pennsylvania
619 Williams Hall
255 S. 36th St.
Philadelphia, PA 19104-6305
joseff@ling.upenn.edu