The Effect of Social Information on the Perception of Sociolinguistic Variables

Nancy Niedzielski

The online version of this article can be found at:

http://jls.sagepub.com/content/18/1/62

Published by:

⑤SAGE

http://www.sagepublications.com

Additional services and information for *Journal of Language and Social Psychology* can be found at:

Email Alerts: http://jls.sagepub.com/cgi/alerts

Subscriptions: http://jls.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://jls.sagepub.com/content/18/1/62.refs.html

>> Version of Record - Mar 1, 1999

What is This?

THE EFFECT OF SOCIAL INFORMATION ON THE PERCEPTION OF SOCIOLINGUISTIC VARIABLES

NANCY NIEDZIELSKI
Panasonic Technologies, Inc.

Forty-one Detroit-area residents were given perceptual tests in which they were asked to choose from a set of resynthesized vowels the tokens that they felt best matched the vowels they heard in the speech of a fellow Detroiter. Half of the respondents were told that the speaker was from Detroit, whereas half were told that she was from Canada. Respondents given the Canadian label chose raised-diphthong tokens as those present in the dialect of the speaker, whereas those given the Michigan label did not. Respondents given the Michigan label chose vowels that were quite different from the Northern Cities Chain-Shifted variety present in the speaker's dialect. Because the "speaker's" perceived nationality was the only aspect that varied between the two groups of respondents, this label alone must have caused the difference in the selection of tokens. This indicates that listeners use social information in speech perception.

Various studies from the fields of acoustic phonetics, sociolinguistics, and social psychology suggest that a variety of different factors are used by listeners in the perception and interpretation of speaker dialect. For instance, researchers have used acoustic phonetics to provide evidence for the effect of several factors that contribute to speech perception: McGurk and MacDonald (1976) showed that visual information about what segment a speaker articulated (i.e., lip movement) affected which segment was perceived; Ladefoged and Broadbent (1957) provided evidence for the fact that hearing as few as three vowels in a speaker's speech influenced the perception of other vowels; Strand and Johnson (1996) showed that "voice" information such as gender influenced the perception of certain fricatives; and Janson (1986) and Willis (1972) showed that dialect background information about a speaker had an influence on the perception of his or her speech. Research in both the fields of sociolinguistics and social psychology has suggested that stereotypes about the social groups speakers are members of (or, are *believed* to be members of) have an influence on how their language varieties are perceived (e.g., Beebe, 1981; Thakerar & Giles, 1981; Williams, 1976). Each of these studies suggests that

62

perception is much more than mere phonetic processing of the speech signal and that other information is used by listeners to process speech signals.

The goal of the study presented here was to determine to what extent listeners use *social information* about a speaker in constructing that speaker's phonological space. The main hypotheses that underlay this research were the following: (a) Listeners use social information just as they use visual and other information to create or calibrate the phonological space of speakers; (b) stereotypes about given language varieties affect the way in which listeners calibrate the phonological space of speakers of those varieties; and (c) people's stereotypes about their *own* variety can be inaccurate, and the phonological space calibrated for members of their own speech communities reflects this inaccuracy.

Previous language-attitudes work in Detroit (Niedzielski 1995, 1997) had yielded several relevant findings for the present research: First, Detroit residents hold stereotypes about the English spoken in Canada, particularly about Canadian Raising (CR). This is significant because CR, a phonetic process whereby the /aw/ diphthong in words such as *house* and *about* are produced with the tongue farther forward and raised higher than it would be for /a/ alone (yielding such caricatures as "aboot the hoose" when, for instance, Detroit residents imitate Canadian English), is found in the English of (White, middle-class) Detroit residents, although these residents do not report any awareness of such raising in their own speech.

In addition, language-attitudes research indicated that Detroiters feel that the dialect of (White, middle-class) Detroit is equivalent to Standard American English (SAE). When questioned directly as to where one is likely to hear standard English spoken, most Detroit residents will confidently answer "right here." However, the dialect used by White, middle-class Detroiters contains several features that are *not* considered to be "standard"; thus, if we accept Wolfram's (1991) description of SAE as a collection of language variants that are not stigmatized (rather than a dialect of any particular region), then the variety used by these Detroiters is not standard. It has been greatly affected by the phenomenon known as the Northern Cities Chain Shift (NCCS), a process that has dramatically shifted the vowels in Michigan, Minnesota, Northern Illinois, Northern Ohio, and Western New York (although its use is most often associated with Minnesota or Chicago), described in greater detail in Labov (1994).

How can it be the case, then, that Detroiters do not notice the CR and the NCCS-vowels in the speech of their fellow Detroiters? Do their beliefs about what Detroit residents *should* sound like have greater influence than the acoustic information that their ears receive? If they believe that a speaker is from an area that is stereotyped for these

features, would they be more likely to notice them? This study is an attempt to begin to answer these questions.

THE EXPERIMENT

Forty-one Detroit-area residents took part in perceptual tests in which they were asked to choose from a set of computer-resynthesized vowels the tokens that they felt best matched the vowels they heard in the speech of a fellow Detroiter (the respondents were told that the author was working for a computer company that was interested in how well computerized vowels matched vowels produced by human speech). About 50 sentences were presented to each respondent aurally, and for each sentence, they were told to concentrate on the vowel they heard in a particular word from that sentence. Next, they were asked to match that vowel to a set of six computer-resynthesized vowels[1] and were told to chose the one that best matched the one that they heard the speaker produce. The vowel space of the speaker that each respondent heard is reproduced as Figure 1.

Half of the respondents were told, however, that the speaker they were hearing was from Detroit, whereas half were told that that same speaker was from Windsor, Ontario, Canada, which is directly across the Detroit River from Detroit. They each heard *the same Detroit speaker*, and because this was the only difference in the two sets of respondents, any differences in the answers from these two sets must be attributed to the expectations that the respondents had based on these two labels.

INFLUENCE OF NATIONALITY LABELS
ON THE PERCEPTION OF CANADIAN RAISING

For the set of words that contain potentially Canadian-raised segments, it was predicted that by leading the Detroiters to believe that a fellow Detroiter was, in fact, Canadian, the Canadian-raised /aw/ would more likely be noticed by the respondents, because raising is a stereotype of Canadian English. The respondents might therefore notice and report hearing a raised variant in the speaker's speech. Conversely, it was predicted that if respondents are told that the speaker was from Detroit, they would instead be less likely to notice the raising and report hearing a nonraised, more "standard" variant, with a nucleus closer to canonical /a/.

It was noted above that respondents had a total of six variants from which to choose the one that best matched the one they felt they heard the speaker produce. In actuality, however, most respondents chose
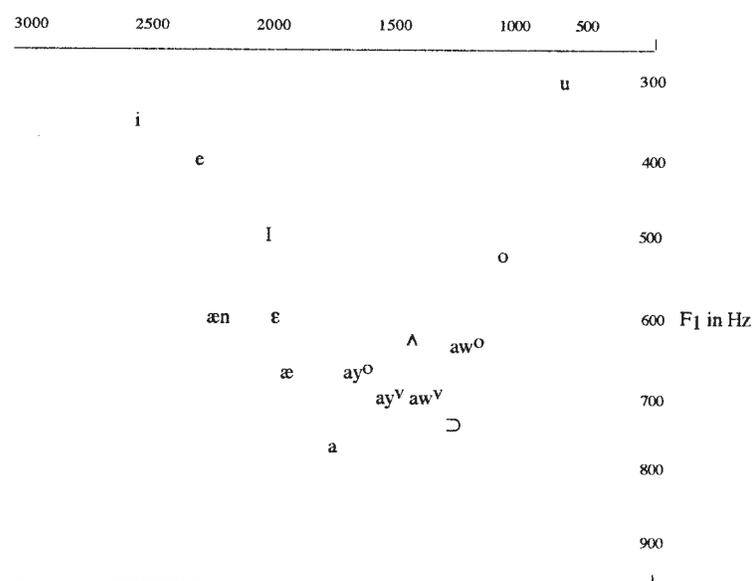
Figure 1.   **Vowel space of the Detroit female speaker on the test tape.**

one of three adjacent variants only.[2] The formant frequencies of these three variants are presented in Table 1. Fourteen outliers (see Note 2) were discarded from the calculations, and all tables and figures contain the results of the effect of perceived nationality on the three tokens that the respondents chose only.[3]

If these values are compared to the speaker's F1 and F2 values found in Figure 1, it is apparent that Token 4 is closest to what the speaker actually produced in these words; for this reason, it is given the label of "actual onset" in this and subsequent figures and tables. This token is raised because the onset of this diphthong is considerably higher than the speaker's /a/. The speaker's actual variant, then, is a *raised* variant.

Token 3 contains a nucleus that is similar to the standard /a/ that is suggested for a female speaker (e.g., in the work of Peterson and Barney, 1952, a standard reference of canonical English vowels), and is therefore labeled as "canonical /a/." This is the variant that the speaker would produce if there was no raising in her dialect.

Token 2 contains an onset that is even lower than a canonical /a/, with an F1 of 900 Hz and an F2 of 1,600 Hz; it is very unlikely that such an onset would be produced in this diphthong (particularly in a prevoiceless environment, in which the duration of the onset would be fairly short), because it moves in the direction away from the back- and upgliding offset /w/ in the diphthong. This token is thus referred to as "ultralow."

Table 1

Formant Values for Tokens Chosen by Respondents for Words Containing /aw/

No. of Token	F1	F2	Label of Onset[a]
2	900	1,600	Ultralow
3	830	1,330	Canonical /a/
4	675	1,150	Actual onset produced by speaker

a. See text for an explanation of these labels.

Table 2

House: *Influence of Nationality Labels on Token Selection*

Token (Label)	2 (Ultralow)	3 (Canonical /a/)	4 (Actual Token)	Total
Canadian	15%	25%	60%	
n	6	10	24	40
Michigan	38%	51%	11%	
n	15	20	4	39

Note. χ^2 = 23.48; *p* < .001.

Table 2 contains the raw data for the percentage of times that a particular token was chosen as the one that respondents felt best matched the speaker's vowel in the word *house*. Figure 2 contains the same data in a graph format.[4] The numbers and percentages refer to the number of times that a particular token was selected by the respondents as the one that best matched the vowel produced by the speaker in that word.

Several things become apparent when looking at these. First, the label given to the respondents regarding the speaker's nationality influenced their perception of these vowels (significant at *p* < .001, based on a chi-square test). Each respondent heard the same female speaker from Detroit, and the test was exactly the same for every respondent, except that the word *MICHIGAN* or *CANADIAN* was written in red across the top of each test page. Therefore, the only thing that could have caused the difference in the respondents' choice of tokens at this statistically significant level was the "Canadian" or "Michigan" label.

Several other important facts emerge from the data contained in Table 2. One significant difference between the two groups of respondents is the percentage of those respondents who chose the actual token. As the table shows, 60% of the respondents who thought the speaker was Canadian chose Token 4, the token closest to the actual variant produced by the speaker. In contrast, only 11% of the respondents who were told that she was from Michigan chose this token. Eighty-nine percent of these respondents, therefore, chose a token that was lower than the actual variant, and almost 40% chose a token that was a great deal lower than *any* /aw/ produced by *any* female speaker
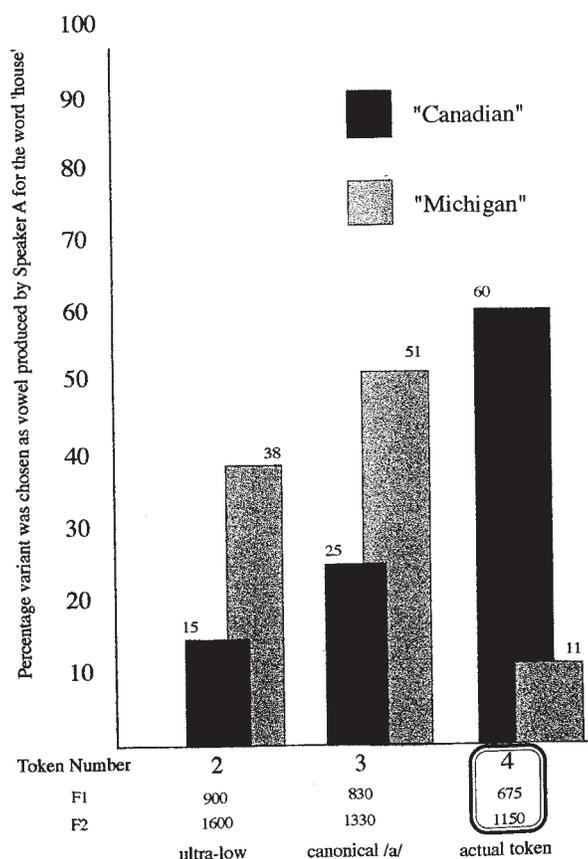
Figure 2. **House**; **influence of nationality labels on selection of tokens.**

in Detroit, even in prevoiced environments—the ultralow one. In contrast, only 15% of the respondents who were told that the woman was Canadian chose this ultralow Token 2 as the one that matched the diphthong that they heard in this word.

Table 3 displays the data from *all* of the words containing /aw/ in the test; Figure 3 contains the same information in graph format. It should be apparent from the table and the graph that the label the speaker was given did, in fact, influence how the respondents perceived this diphthong in each of these words. These tables reveal that those respondents who thought the speaker was Canadian reported hearing a raised token, which the Detroit speaker did produce. If the respondents were told (correctly) that the speaker was from Michigan, they misidentified the diphthong as one much lower than she actually produced.

Table 3
Words Containing /aw/: Influence of Nationality Labels for All Respondents

Token (Label)	2 (Ultralow)	3 (Canonical /a/)	4 (Actual Token)	Total
Canadian	15%	32%	53%	
n	23	50	83	156
Michigan	40%	45%	15%	
n	65	74	25	164

Note. See text for formant values of tokens. $\chi^2 = 54.26$; $p < .001$.
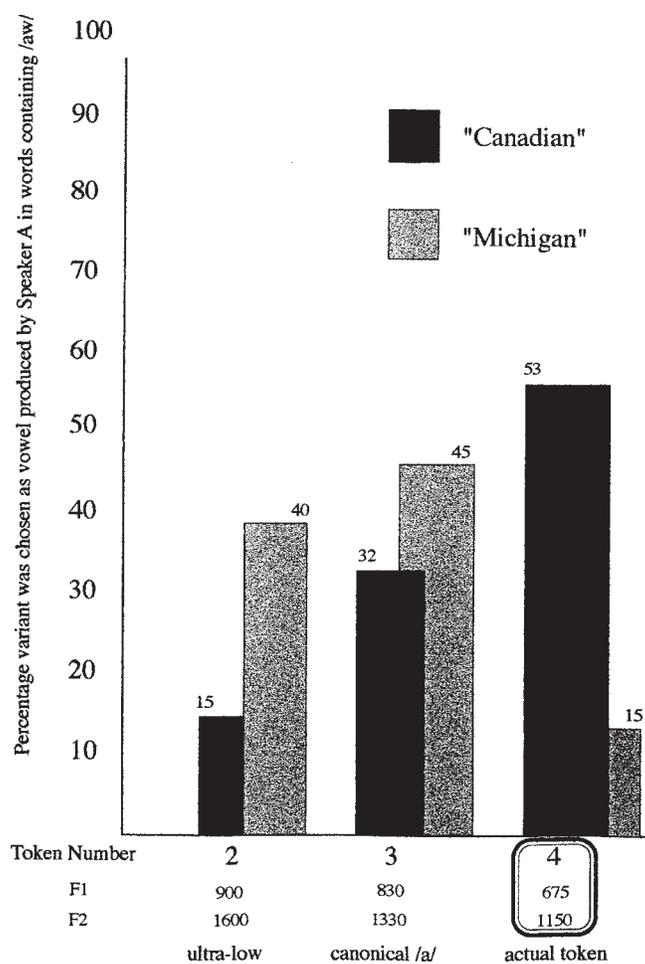

Figure 3. **All words containing /aw/; influence of nationality labels on selection of tokens.**

These findings provide evidence for the conclusion that CR is a Canadian stereotype for Detroit residents, while remaining a virtually unnoticed feature of their own dialect.

Gender is an interesting variable with regard to CR. Niedzielski (1997) reports on the differences between reporting CR as a stereotype of Canadian English, with the primary difference being that the female respondents were much more likely to offer raised diphthongs as an example of a feature that separates Michigan from Canadian English. However, no gender difference was present in this study. As Table 4 shows, there was essentially no difference between what male and female respondents selected in either the "Canadian" group or the "Michigan" group. Thus, whereas female Detroiters overtly state that CR is a feature of Canadian English, and behave accordingly in the perceptual test, male respondents are less likely to report that this is a feature of Canadian English, yet they behave on the perceptual test as if this is a feature that they assign to Canadian speakers. For both genders, the nationality label that was given to the speaker influenced the respondents' choices of tokens to the same degree. Both women (who reported this stereotype) and men (who did not) were more likely to choose the actual, raised variant if they were told that the speaker was Canadian and the lowered tokens if they were told that she was from Michigan. This is discussed further below.

This section provides evidence for the fact that for words that contain the stereotyped /aw/,[5] listeners "hear" the stereotyped raised variant if the speaker fits the social description of someone who is expected to raise it—that is, someone from Canada. If, however, the speaker does *not* fit this social description—if the speaker is believed to be from Michigan—then listeners are less likely to "hear" or notice the raised variant.

INFLUENCE OF NATIONALITY LABELS ON CHAIN-SHIFTED WORDS

Several test segments also were run with words that contained vowels that are part of the NCCS. The effect of the labels "Canadian" versus "Michigan" on the perception of these vowels was also investigated. As in the case of the perception of /aw/, gender was not a significant factor, and thus the tables below display the combined data of men's and women's results.

Table 5 presents the results of the labels "Canadian" versus "Michigan" on the perception of the vowel /a/ in *pop*, and Figure 4 displays the same results in graph format. Although the results of the effect of *nationality labels* are not statistically significant, there are other more provocative findings in Table 5. The formant frequency values for the three tokens that respondents selected are displayed in Table 6.

Table 4

Words Containing /aw/: Influence of Nationality Labels by Gender

Token (Label)	2 (Ultralow)	3 (Standard /a/)	4 (Actual Token)	Total
Canadian				
Female	16%	31%	53%	
n	13	25	42	80
Male	13%	33%	54%	
n	10	25	41	76
Michigan				
Female	41%	45%	14%	
n	32	35	11	78
Male	38%	45%	16%	
n	33	39	14	86

Table 5

Pop: *Influence of Nationality Labels on Token Selection*

Token (Label)	1 (Hyperstandard)	2 (Canonical /a/)	3 (Actual Token)	Total
Canadian	3%	93%	5%	
n	1	37	2	40
Michigan	10%	85%	5%	
n	4	36	2	42

Note. The results of the effect of nationality labels are not statistically significant.

First, almost none of the respondents in either group chose the actual variant that the speaker produced, which was Token 3. This vowel is quite fronted and raised, because of the NCCS. The formant values for this vowel are 700 Hz for F1 and 1,600 Hz, which is fairly close to a canonical /æ/ (several California residents who participated in the pilot study felt that there must be an error—this vowel could not be perceived as an /a/; to them, it sounded like *pap*). Just 5% of the Detroit respondents in either group chose this token.

The respondents in both groups overwhelmingly chose Token 2 instead. This token had an F1 of 900 Hz and an F2 of 1,400 Hz, which is close to a more standard /a/ (as suggested for female speakers in Peterson & Barney, 1952). The vowel that the respondents thus selected is a vowel that conforms to a standard or canonical /a/, rather than the vowel that the speaker actually produced.

In addition, 10% of the choices made by respondents who were told that the speaker was from Michigan were for Token 1. This token was close to a canonical /ɔ/, so that F1 was 770 Hz and F2 was 1,050 Hz, and makes the word sound like *paup* instead of *pop*. It is a "hyperstandard" version of /a/, not necessarily a standard one. Examining the speaker's
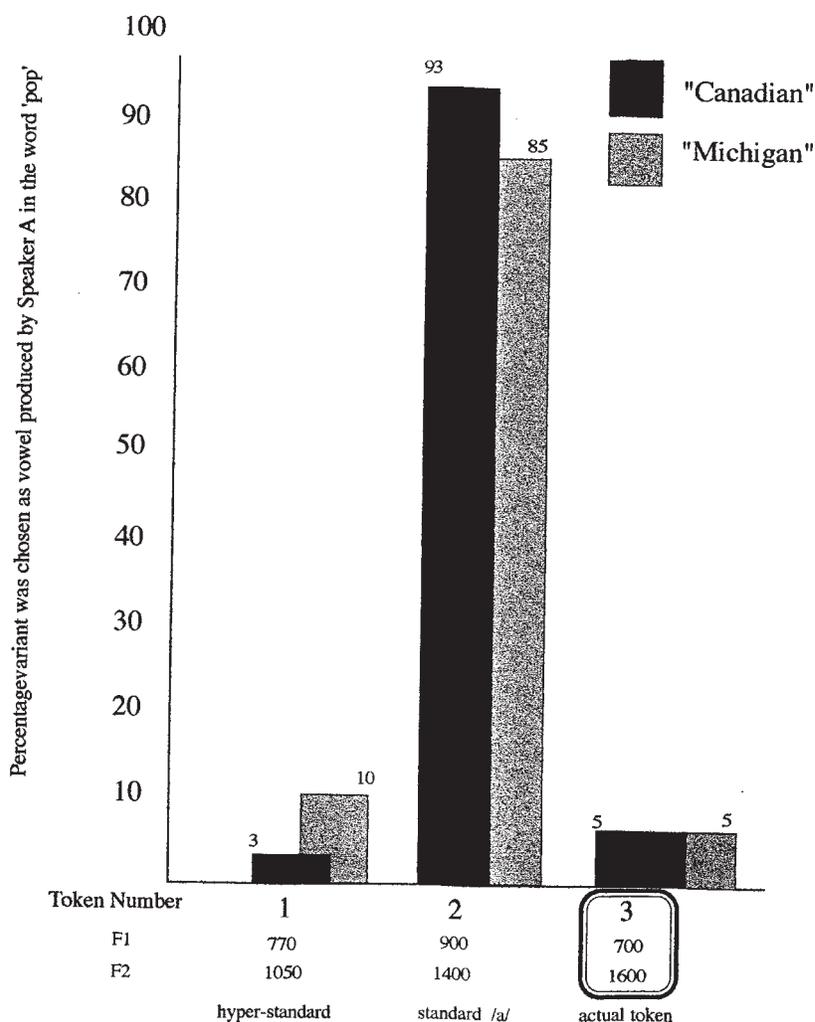
Figure 4. ***Pop***; **influence of nationality labels on selection of tokens.**

vowel space in Figure 1 reveals why this token is characterized as hyperstandard: The movement from the actual variant that the speaker produced as /æ/ to a standard /a/ is toward the back of the vowel space. Although this hyperstandard variant is produced farther back than the actual variant, it is actually farther back in the vowel space than the standard /a/; like the standard variant, F2 is lower than the F2 of the actual variant. It is labeled *hyperstandard*, however, because F2 is even lower for this token. It is significant that this token was not chosen to the same extent by respondents who thought that

Table 6
Formant Values of Tokens Chosen by Respondents for Pop

No. of Token	F1	F2	Label of Token[a]
1	770	1,050	Hyperstandard
2	900	1,400	Standard /a/
3	700	1,600	Actual vowel produced by speaker

a. See text for an explanation of these labels.

Table 7
Last: *Influence of Nationality Labels on Token Selection*

Token (Label)	1 (Hyperstandard)	2 (Standard /æ/)	3 (Actual Token)	Total
Canadian	3%	87%	10%	
n	1	34	4	39
Michigan	10%	90%		
n	4	38	0	42

Note. The results of the effect of nationality labels are not statistically significant.

the speaker was Canadian; instead, 93% of these respondents chose the more standard, canonical variant. This will be discussed further in the following section.

Again, the fact that these results are not significant according to a chi-square test is not necessarily relevant; the main significance for these NCCS words is that respondents are not perceiving the variants that the speaker actually produced. This is discussed further in the following section.

The same pattern holds for the vowel /æ/ in the word *last*, displayed in Table 7 and Figure 5. The formant frequency information for the three tokens that the respondents selected is shown in Table 8.

Not one of the respondents who were told that the speaker was from Michigan chose Token 3, which was the variant the speaker actually produced. This token had an F1 of 700 Hz and an F2 of 1,900 Hz, demonstrating a raised and fronted variant that is expected from the NCCS (making the word sound similar to *lee-ist* to non-NCCS speakers). Instead, respondents in both categories chose the lower, more central Token 2, which had an F1 of 775 Hz and an F2 of 1,700 Hz. This is much closer to a "standard" /æ/ (as suggested for female speakers in Peterson & Barney, 1952).

Once again, however, certain respondents (10%) who were told that the speaker was from Michigan chose a token that could be considered "hyperstandard," for the same reasons as above: Although this token moves in the same direction toward the standard (i.e., F1 is higher than for the actual variant), it actually moves *past* canonical /æ/ and
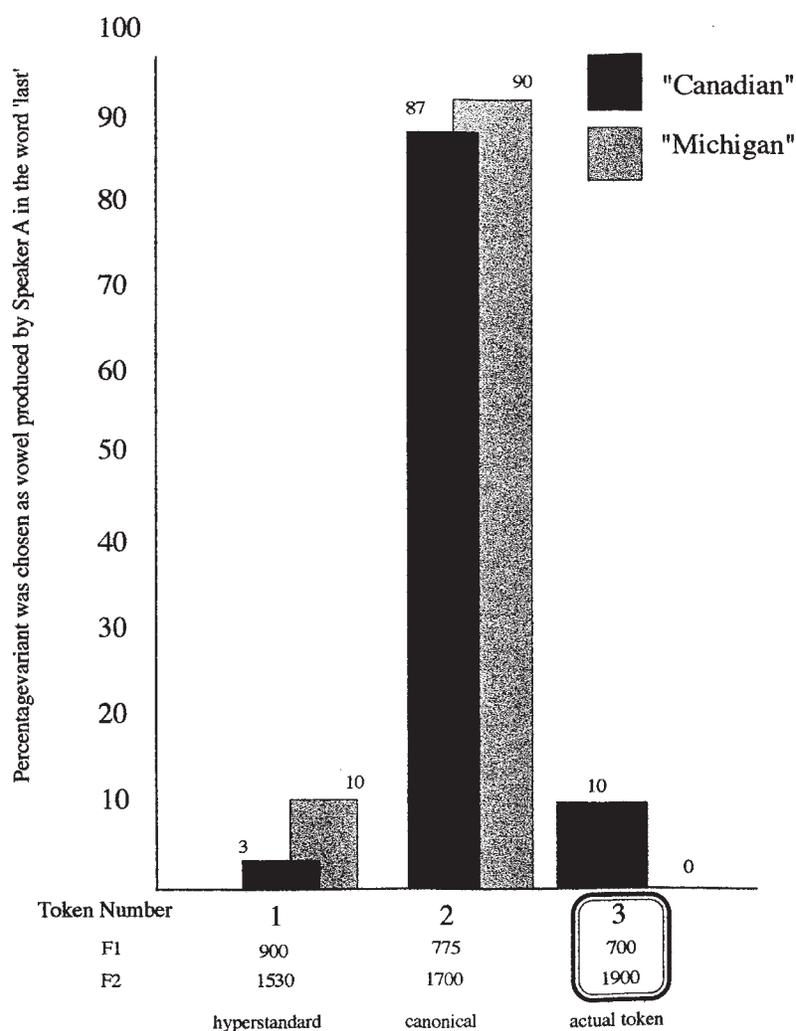
Figure 5. **Last; influence of nationality labels on selection of tokens.**

approaches a canonical /a/ (so that it would sound similar to *lost*), with an F1 of 900 Hz and an F2 of 1,530 Hz. Just as for /a/ (described above), this token is thus considered "hyperstandard." Once again, it was not chosen by as many of those respondents in the Canadian group, who for the most part chose either the more "standard" Token 2, whereas 10% chose the actual Token 3.

Table 9 displays the results of the perception of the vowel /o/ in the word *close*. The formant values are presented in Table 10. Although this word was only presented a single time and thus yielded a small

Table 8
Formant Values of Tokens Chosen by Respondents for Last

No. of Token	F1	F2	Label of Vowel[a]
1	900	1,530	Hyperstandard
2	775	1,700	Canonical /æ/
3	700	1,900	Actual token produced by speaker

a. See text for an explanation of these labels.

Table 9
Close: *Influence of Nationality Labels on Token Selection*

Token (Label)	2 (Standard /o/)	3 (Actual Token)	4 (Standard /U/)	Total
Canadian	40%	40%	20%	
n	8	8	4	20
Michigan	67%	29%	5%	
n	14	6	1	21

Note. The results of the effect of nationality labels are not statistically significant.

Table 10
Formant Values of Tokens Selected by Respondents for Close

No. of Token	F1	F2	Label of Vowel[a]
2	650	1,200	Standard /o/
3	560	1,050	Actual token that the speaker produced
4	500	950	Standard /U/

a. See text for an explanation of these labels.

sample (so that the results are only suggestive at this point), they are provocative enough to warrant examination.

Although almost 30% of the respondents given the Michigan label did choose the actual variant, most chose Token 2, a token that was lowered. The speaker's /o/ is quite high, a realization perhaps attributable to the effect of the NCCS: F1 of her actual variant (Token 3) is 560 Hz, whereas F2 is 1,050 Hz (it may sound similar to *close* to non-NCCS speakers). The token that these respondents chose, however, is closer to the standard /o/ for women presented in Peterson and Barney (1952): F1 is 650 Hz and F2 is 1,200 Hz.

An equal percentage of respondents given the Canadian label chose each of these two tokens. Twenty percent, however, chose Token 4, which is canonical /U/: F1 is 500 Hz and F2 is /800/.

This is interesting because the raised variant of /o/—that is, Variant 3, the one that this Detroit speaker actually produced— was mentioned by several of the Michigan respondents in the language-attitudes survey as a feature of a language variety associated with Minnesota (usually respondents produced the word *Minnesota* with a raised /o/ as an illustration of this). This variable, like the Canadian Raising presented in the Experiment section, is thus a feature present in Detroit speech but stereotyped for a different geographical region. Further investigation with a larger sample is needed before the similarity to Canadian Raising is concluded, and this is discussed below.

Table 11 displays the effects of nationality labels on the perception of the vowel in the word *bed*.[6] The formant values of the tokens are presented in Table 12. Token 1 is the variant that the speaker actually produced (and because she is from Detroit and this vowel is part of the NCCS, her /e/ is quite central). F1 for this vowel is 600 Hz, and F2 is 1,700 Hz. However, as Table 11 shows, almost no one chose this token. Instead, the respondents given the Michigan label chose a token quite similar to canonical /ɛ/ (see Peterson & Barney, 1952). F1 was 450 Hz, and F2 was 1,950 Hz for this token.

Nevertheless, the respondents who were told that the speaker was from Canada were more likely to choose fronted and lowered Token 4, similar to canonical /æ/. The F1 value of this token was 700 Hz, whereas the F2 value was 2,000 Hz. This means that respondents in the Canadian group were attributing a pronunciation similar to *bad* (/bæd/) to the speaker.

Interestingly, this is not too far from what Canadians in Ontario are doing with this vowel (Meechan, 1996). In several varieties of Canadian English, /ɛ/ is lowering to /æ/ (but not centralizing to /ʌ/, as in the NCCS), and so these respondents are inadvertently assigning a variant similar to one that an Ontario speaker would produce, to the Detroit speaker.

While the respondents in the Canadian group were not as accurate at choosing the actual variant that the speaker produced, they did succeed in choosing a variant that is similar to what Ontarians are using. It is not clear, though, that this is due to Detroiters' knowledge of Canadian speech. Instead, it may be due to the tendency that evidenced in the previous cases of these NCCS-vowels: The respondents in the Canadian group assign a nonstandard pronunciation to the speaker because they think she is a Canadian, whereas those who are told that the speaker is from Michigan assign a more standard pronunciation to her. As we have seen, very few of the respondents in the Michigan group chose the actual variant that the speaker used.

This seems to suggest that when a respondent is presented with a speaker that he or she thinks is a fellow Michigander, the stereotype of Michigan speakers as speakers of "standard" English emerges, and the respondent thus selects a standardized, canonical set of vowels.

Table 11
Bed: *Influence of Nationality Labels on Token Selection*

Token (Label)	1 (Actual Token)	2 (Canonical /ɛ/)	3 (Canonical æ)	Total
Canadian	0%	30%	70%	
n	0	6	14	20
Michigan	5%	67%	28%	
n	1	14	6	21

Note. $\chi^2 = 7.38$; $p < .05$.

Table 12
Formant Values of Tokens Chosen by Respondents for Bed

No. of Token	F1	F2	Label of Onset[a]
1	600	1,700	Actual variant produced by the speaker
2	450	1,950	Standard /ɛ/
3	700	2,000	Standard /æ/

a. See text for an explanation of these labels.

This becomes quite clear in viewing the data presented in Table 13 (and reproduced in graph format in Figure 6). This table contains the effects of nationality labels and the selection of tokens for the vowel /ɔ/ in the word *talking*. This vowel is shifting dramatically in North American English, so that it has merged with the vowel /a/ (making *caught* and *cot* homonymous) in several dialects, including southern California, Canada, and some northern cities.

As Wolfram (1991) has pointed out, however, the feelings about the merged /ɔ/ versus the unmerged /ɔ/ are ambiguous, in terms of which is more standard. Therefore, although most Americans would choose the unmerged variant as more standard, it is quite possible to hear the merged variant used by, for instance, the national media.

In addition, Preston (1987) shows that there is ambiguity in self-reports of this variable in a different Michigan city. In this work, Preston reports the results of a survey of about 1,500 native Michiganders,[7] who performed a rhyming task. Respondents were asked to match several words (which contained vowels involved in the NCCS) with words that they felt had similar-sounding vowels. For instance, a respondent might place the word *miss* together with the words *pit* and *sick*. He found that for "European Americans," (i.e., White respondents), there was considerable variability into which group such words as *dog* and *clawed* (both of which contain /ɔ/ in Michigan dialects) were put. He found that although most respondents matched these words up with words like *moth* and *caught* (which both also contain /ɔ/), almost 20% of the respondents matched the words *dog* and *clawed* with *clock* and *slob* (which contain an /a/ in standard dialects). In addition, 10% of

Table 13
Talking: *Influence of Nationality Labels on Token Selection*

Token (Label)	1 (Hyperstandard)	2 (Actual Token /ɔ/)	3 (Merged Token)	Total
Canadian	0%	15%	85%	
n	0	3	17	20
Michigan	9%	29%	57%	
n	2	6	12	21

Note. The results of the effect of nationality labels are not statistically significant.

the respondents in this study matched words containing /a/, such as *not* and *Don*, with words containing /ɔ/.

Thus, even in dialects that are not merging /ɔ/ and /a/, there is ambiguity about which variants—the unmerged pair or the merged single variant—is more "standard." If we continue to appeal to Wolfram's (1991) view of a "standard" feature as one that is not marked for ethnicity, region, social class, and so forth, then both the merged and the unmerged variant can be considered standard. This ambiguity receives further support in the present study.

The pattern seen here follows a very different pattern from that seen so far for the respondents in the Michigan group. Fifty-seven percent of the respondents who were told that the speaker was from Michigan chose Token 3—a token that represents a canonical /a/. This token had an F1 of 880 Hz and an F2 of 1,500 Hz. (Interestingly, this /a/ is not even the /a/ that the speaker produced—hers was higher and farther front [see above]; rather, these respondents chose a standardized version.) The selection of this token is consistent with the other variables examined from the NCCS, *if* the merged variant /a/ is standard in this word. However, this is the lowest percentage for the selection of the "standard" variant of any of these NCCS variables. This may reflect that fact that the standard for words containing /ɔ/ is ambiguous.

In addition, almost 30% of these respondents chose the actual variant (Token 2), which is close to a standard /ɔ/: F1 is 720 and F2 is 1,300, the highest percentage for any of these NCCS variables. Furthermore, almost 10% chose a hyperstandard token that approached /o/, with an F1 of 700 Hz and an F2 of 1,100 Hz.

This Michigan group was thus split 57%/40% between a variant that is found in merged dialects and a variant that is found in unmerged dialects. This seems to suggest that although the respondents assign standard variants to this Michigan speaker, the "standard" for this vowel is not so clear.

The respondents in the group that was told the speaker was Canadian predominantly chose the merged /a/ token, and only 15% chose /ɔ/ (and none chose the hyperstandard token). This is puzzling: If the
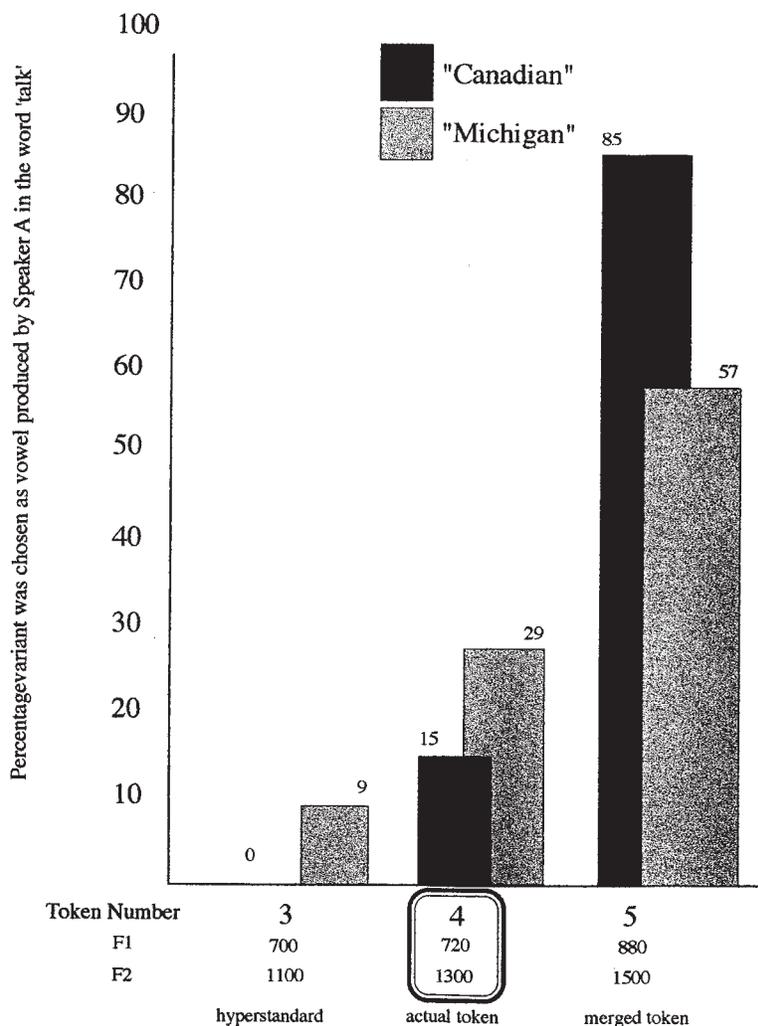
Figure 6.   ***Talking***; **influence of nationality labels on selection of tokens.**

merged variant is beginning to be viewed as "standard," and the selection of this variant by those who were given the Michigan label seems to suggest this is the case, then why is this token more often selected by those given the Canadian label? This is the first variable for which the more standard variant was assigned by this group of respondents.

The fact that this variable does not appear to pattern with the other variables in this study provides further evidence for the fact that no clear standard for this variable has emerged.

DISCUSSION

ON CANADIAN RAISING

As the previous section clearly shows, social information about the speaker *does* have a significant effect on speech perception, even at the phonetic level. As the language-attitudes work shows, Detroit speakers do not hear CR in their own speech and in the speech of their Detroit neighbors. However, it is possible to make the CR in the speech of Detroiters noticed by the speakers themselves—by leading them to think that the speaker is Canadian. Detroiters expect to hear raised /aw/ in the speech of Canadians, and therefore, they do. They do not expect to hear it in the speech of fellow Detroiters, and therefore, they do not.

In addition, if CR is pointed out in a Detroiter's speech, that Detroiter is said to sound "Canadian." There is some evidence that this feature may just now be beginning to be noticed in the speech of some Detroiters by Detroit residents. A Detroit resident pointed out to me when I was in Detroit that one radio announcer sounded Canadian to her. When I listened to this announcer, she sounded to me like a typical (White middle-class) Detroiter. Several weeks later,[8] this same radio announcer was being teased by her coworkers for sounding Canadian. She protested that she had only been in Canada four times in her life and was born and raised in Troy, Michigan, a suburb of Detroit, but her coworkers insisted that she sounded Canadian (although they did not provide any specific reasons for this). The fact that CR is just starting to be noticed may mean that in time, it will become a self-stereotype of Detroit speech.

The question of gender differences and stereotypes remains, however. Why is it that gender did not play as significant a role in this perceptual study as it did in the language-attitudes study? In other words, why did so many men report that there was no difference between the English of Michiganders and the English of Canadians, and yet, so many men select tokens differently according to those very same labels?

Part of the answer may be that the men in the language-attitudes study were less willing to share stereotypes about other speakers with me, because I am a woman, and hence, I may have been regarded as an "out-group" member. Evidence for this can be found in the fact that women were much more likely to give me detailed and elaborate answers, often illustrating their beliefs with personal narratives. Men, on the other hand, were more likely to provide one-sentence answers to the questions and were not as likely to provide personal narratives. Thus, they may have the same stereotypes of Canadian speakers as the women, although they may have simply been less willing to share them with me.

One group of men appears to provide counterevidence to the above explanation, and this is the group of hockey fans. Every man who mentioned that he was a fan of hockey stated that there were differences between Canadian and Michigan English. These men often did provide detailed descriptions of the differences, which for them included stress patterns, lexical items, and even word pronunciations.

Hockey is quite strongly associated with Canada. It is the only professional sport that has so many Canadian teams that compete with American teams and is seen at least as equally "Canadian" and "American" (as opposed to basketball and baseball, in which only a few Canadian teams compete and which are seen as "American" sports). It may be the case that for these men, the distinction between Canadian and American was stronger than the out-group distinction of male and female, and that this stronger distinction overcame the reticence that the other nonhockey fans felt in discussing their stereotypes.

A more likely explanation is that the hockey fans' greater familiarity with Canadian English may have allowed them to speak with greater confidence about differences than men who had not had as much exposure to it. Perhaps the fact that I was an out-group member meant that men were only willing to share ideas that they had more concrete evidence for. The hockey fans may have felt that they had more "evidence" for Canadian differences than the nonhockey fans did.

Interestingly, it was a hockey fan in the present study who had the most difficult time with the task. This respondent flatly refused to believe that the speaker on the tape was Canadian and told me he chose Michigan vowels because the speaker was clearly from Michigan; in fact, he did pattern with the respondents given the Michigan label. His more intimate knowledge of a Canadian dialect, which may have included those items mentioned above (such as stress patterns and lexical items), caused him to reject the label that the others accepted.

My conclusion regarding gender is that most Detroit men do have, at least at some level, the same stereotypes of Canadian speech as Detroit women, and the results of the perceptual study certainly seem to reveal this. I was unable to discover these stereotypes in the language-attitudes study, however, because the male respondents may not have felt comfortable enough with me to reveal stereotypes. Of course, an additional attitudes study using a male interviewer is needed to positively conclude this.

ON THE NCCS

The results of this portion of the perceptual study provide further evidence that White speakers in Detroit feel that they are speakers of "standard" speech. Several previous studies have shown that Michigan speakers display a high degree of linguistic security (cf. Preston,

1989). To date, however, no other study sought to determine whether Detroiters felt that the Northern Cities Chain Shifted vowels were "correct" and standard or whether they simply did not hear the shift in their own speech.

The present study provides strong evidence for the latter. Even when faced with acoustic data that suggest otherwise, Detroit respondents select standard vowels as those that match the vowels in the speech of fellow Detroiters. It is not the case, then, that Detroiters assign standard labels to raised peripheral vowels and lowered lax ones, that is, NCCS vowels. Rather, Detroiters simply do not perceive NCCS vowels at this level of consciousness.

Figure 7 demonstrates this. In this figure, the vowels that are bolded and enlarged represent the vowels that most respondents given the Michigan label chose (overlaid on the speaker's vowel space from Figure 1). What emerges is a vowel chart that looks quite close to the vowel space created from standard formant values of Peterson and Barney (1952) for female speakers. The respondents assigned a vowel space to the speaker that represents what most speakers of English would call "standard" and, when compared to the speaker's actual values, reveal how far from the actual vowels the perceived vowels are.

One important question to ask, then, is where this idea of a standard vowel system comes from. The first possibility is that the correlation of SAE with the language varieties in Michigan may have a historical basis. In Labov, Yaeger, and Steiner (1972), those respondents who were older than 50 did not show the same degree of the NCCS as the younger speakers; /æ/ was not raised at all in most of the speakers' vowel spaces, and this is thought to be the *first* step in the NCCS (such age stratification provides evidence that the NCCS is a change in progress). One hypothesis for why Michigan speakers regard their own dialect as standard is that previous generations of Detroit speakers used a language variety that was close to standard, and it is this variety that provides their reference point for what "standard" English sounds like.

The problem with this hypothesis is that the research for the Labov et al. (1972) study was performed more than 25 years ago, so that the oldest speakers would now be over 80 years of age. The question becomes, How do later generations pass on the notion that an earlier generation's language variety is standard, when this later generation does not use the variants of the earlier generation? Do people really look at several generations above them, in deciding which language variety is standard? If this was the case for the phenomenon seen in Michigan, it would mean that a younger speaker determines that an older speaker's speech is standard, then uses this standard as the basis for other perceived standard speakers. Until there is evidence that listeners judge older Detroit speakers' language varieties as more standard than younger speakers' varieties, this hypothesis cannot be accepted.
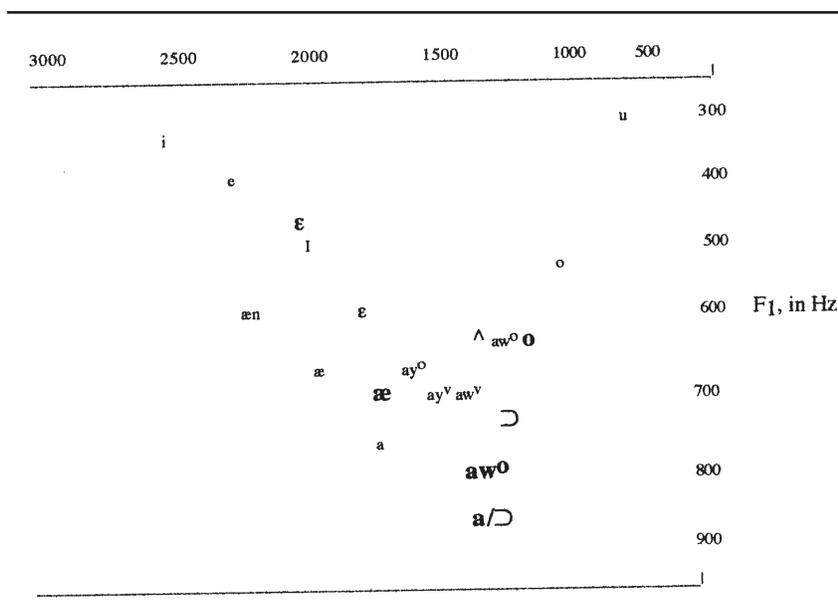
Figure 7.   **Vowel space chosen for speaker on test tape (in bold), superimposed on her actual vowel space (as in Figure 1).**

An alternative hypothesis is that it is the mass media industry that sets the standard. It is possible that because most of the films and television shows in the United States are produced in southern California, the dialect of this region is what SAE is based on. There has been very little acoustic analysis of the language varieties in southern California, however, so that this claim cannot be substantiated. Furthermore, this proposal would not explain why Michigan varieties are also stereotyped as standard. In addition, strong stereotypes about southern California exist in Detroit, so that this dialect does not seem to be one that Detroiters consider to be standard. It seems unlikely, therefore, that this dialect would then serve as a model of SAE for these Detroit residents.

Education, too, may play a part in the speakers' ideas about SAE. The current debate about Ebonics has clearly demonstrated that educators are very concerned with teaching their students to use standard English. However, this standard English is usually restricted to rules of syntax and lexicon. There have been studies that have shown that phonology is relevant in how teachers evaluate students, but the phonological variants that teachers notice are generally highly stigmatized ones, such as those used by some African American inner-city students.

A fourth hypothesis is based on a proposal that Wolfram (1991) has advanced. He suggests that any language variants that are not stigmatized are considered to be standard. SAE may be a conglomeration of

all such nonstigmatized features. The problem with this proposal, however, is that certain features of Detroit varieties are stigmatized, for example, the raised /æ/ in words such as *man*, the fronted /a/ in words such as *father*, and the raised /o/ in words such as *phone*—*if* they are noticed. These features are usually associated with regions other than Michigan, such as Minnesota or Wisconsin. Several respondents in the language-attitudes survey imitated the raised /o/ in the pronunciation of *Minnesota*. The language varieties of Minnesota and Wisconsin are quite similar to the language varieties in Michigan, and each of the speakers whose speech was acoustically analyzed for this study showed such a raised /o/ in his or her speech. Although it is not clear exactly why the dialect of White, middle-class Michigan is not stigmatized in Michigan, whereas the dialect of White, middle-class Minnesota is. This implies that the nonstandard features are noticed if the speaker is believed to be from Minnesota, whereas standard features are perceived if the speaker is believed to be from Michigan, a hypothesis that will be tested by administering a similar speech perception experiment with the labels *Michigan* and *Minnesota*.

A final hypothesis about Michigan residents' self-stereotype as standard speakers is that believing that one speaks SAE is what is most "comfortable" to believe. Perhaps there is a natural human tendency for people to want to belong to the normative group and to believe that they are not "different" (cf. Giles, 1973). Until speakers are made aware of the fact that they are "different," perhaps they do not realize this. This hypothesis, too, is consistent with Wolfram's (1991) idea that until a dialect is stigmatized, it is viewed as standard, perhaps most of all by the speakers themselves.

Furthermore, it may take a critical mass of people stereotyping a variety as nonstandard before the nonstandard stereotype becomes a part of cultural knowledge, both in general and for the speakers of the variety in question.

Perhaps it is a combination of the above factors that led Detroit residents to perceive basically SAE features in the speech of their relatives and neighbors. In any case, this essentially cross-disciplinary study, which used insights from the fields of acoustic phonetics, sociolinguistics, and social psychology, has provided evidence that social information is used by listeners to calibrate the phonological space of speakers. In addition, it provides evidence that stereotypes about given language varieties affect the way in which listeners calibrate the phonological space of speakers of those varieties, that people's stereotypes about their own variety can be inaccurate, and that the phonological space calibrated for members of their own speech communities therefore reflects this inaccuracy. The findings presented above seem to support the basic hypotheses:

1. Listeners do use social information to calibrate the phonological space of speakers.
2. Stereotypes about given language varieties do affect the way in which listeners calibrate the phonological space of speakers of those varieties.
3. People's stereotypes about their own variety are inaccurate, and the phonological space calibrated for members of their own speech communities reflects this inaccuracy.

The results of this study suggest social information must be included in future research in phonetics, sociolinguistics, and social psychology, particularly in the areas of speech perception and language change.

NOTES

1. The vowels were resynthesized from the speaker's actual speech, so they sounded fairly natural; the respondents were given several practice sets, so as to get used to the task.

2. Fourteen out of 758 total choices, or less than 2% of the total choices made by all respondents for each word, were outliers. These 14 outliers were discarded from the calculations. In nine of these cases, the choice of an outlier was made at the end of the test, for a word that the respondents had already matched, although the tokens were presented in reverse order. In these cases, the respondent circled the same number as he or she had the first time, making it possible that the respondent merely copied his or her first answer, without listening to the tokens.

3. For the complete list of formant values, see Niedzielski (1997).

4. There were 20 respondents who were told that the speaker was from Canada and 21 respondents who were told that she was from Michigan. This means that there should be at least 20 tokens for the "Canadian" label and 21 for the "Michigan" label if the word was presented once and 40 for the Canadian label and 42 for the Michigan label if the word was presented twice. There are some figures in the Total column that do not add up to these numbers. The reason for this is that the outliers were excluded; see above.

5. The word *night*, which contains a different diphthong that is also part of Canadian Raising, was also tested in this study; the results of this test seem to suggest that the nationality label did not have as great an effect on the selection of the variant, perhaps because the raised /ay/ is not a stereotype of Canadian English. However, there were several problems with this test segment, so the results are not reported in this study.

6. Recall that only the variants most often chosen by the respondents are displayed here.

7. These respondents were students at Michigan State University in East Lansing, Michigan.

8. Radio station WDRQ, September 4, 1996; 11:00 a.m.

REFERENCES

Beebe, L. (1981). Social and situational factors affecting communicative strategy of dialect code-switching. *International Journal of the Sociology of Language*, *32*, 139-149.
Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, *15*, 87-105.

Janson, T. (1986). Sound change in perception: An experiment. In J. Ohala & J. Jaeger (Eds.), *Experimental phonology* (pp. 253-260). Orlando, FL: Academic Press.

Labov, W. (1994). *The principles of language change*. New York: Blackwell.

Labov, W., Yaeger, M., & Steiner, R. (1972). *A quantitative study of sound change in progress*. Philadelphia: U.S. Regional Survey.

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, *29*(1), 98-104.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*, 746-748.

Meechan, M. (1996, October). *One of us says milk and the other says melk: Lax vowel lowering in Canadian English*. Paper presented at New Ways of Analyzing Variation 25, Las Vegas, NV.

Niedzielski, N. (1995). Acoustic analysis and language attitudes in Detroit. In M. Meyerhoff (Ed.), *(N)Waves and means: University of Pennsylvania working papers in lingusitics* (Vol. 3, pp. 73-86). Philadelphia: University of Pennsylvania Press.

Niedzielski, N. (1997). *The effect of social information on the phonetic perception of sociolinguistic variables*. Unpublished doctoral dissertation, University of California, Santa Barbara.

Peterson, G., & Barney, H. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24*, 2175-2184.

Preston, D. (1989). *Perceptual dialectology*. Dordrecht, the Netherlands: Foris.

Preston, D. (1997). The Northern Cities Chain Shift in your mind. In A. Thomas (Ed.), *Issues and methods in dialectology* (pp. 37-45). Bangor, Wales: University of Wales, Department of Linguistics.

Strand, E., & Johnson, K. (1996). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural language processing and speech technology: Results of the 3rd KONVENS Conference, Beilefeld, October, 1996* (pp.318-336). Berlin: Mouton.

Thakerar, J., & Giles, H. (1981). They are—so they speak: Noncontent speech stereotypes. *Language and Communication*, *1*, 251-256.

Williams, F. (1976). *Explorations of the linguistic attitudes of teachers*. Rowley, MA: Newbury House.

Willis, C. (1972). Perception of vowel phonemes in Fort Erie, Ontario, Canada, and Buffalo, New York: An application of synthetic vowel categorization tests to dialectology. *Journal of Speech and Hearing Research*, *15*, 246-255.

Wolfram, W. (1991). *Dialects and American English*. Englewood Cliffs, NJ: Prentice Hall.