PERCEPTUAL AND PHONETIC EXPERIMENTS ON AMERICAN ENGLISH DIALECT IDENTIFICATION

THOMAS PURNELL
University of Wisconsin–Madison

WILLIAM IDSARDI
University of Delaware

JOHN BAUGH
Stanford University

The ability to discern the use of a nonstandard dialect is often enough information to also determine the speaker's ethnicity, and speakers may consequently suffer discrimination based on their speech. This article, detailing four experiments, shows that housing discrimination based solely on telephone conversations occurs, dialect identification is possible using the word hello, and phonetic correlates of dialect can be discovered. In one experiment, a series of telephone surveys was conducted; housing was requested from the same landlord during a short time period using standard and nonstandard dialects. The results demonstrate that landlords discriminate against prospective tenants on the basis of the sound of their voice during telephone conversations. Another experiment was conducted with untrained participants to confirm this ability; listeners identified the dialects significantly better than chance. Phonetic analysis reveals that phonetic variables potentially distinguish the dialects.

Nationally, African Americans lag behind Whites in median home value (61%; U.S. Census Bureau, 1995). This disparity may follow from such complex and interrelated issues as lower-paying employment (as indicated by a median family income 58% of White households), self-segregation, and housing discrimination. Varying levels of influence have been attributed to discrimination in the segregation of neighborhoods, from a prominent role where discrimination accounts for 20% to 30% of segregation (Courant, 1978; Tobin, 1982; Yinger, 1986) to a minor role of no more than 15% (Clark, 1986, 1993; Galster & Keeney, 1988). Myers and Chan (1995) place race as a reliable indicator of discrimination in mortgage lending, accounting for 70% of the gap in

rejection rates while controlling for other factors. Regardless of the size of the effect, discrimination is indisputably present in our society and led Congress in 1988 to amend laws protecting against housing discrimination.

Intentional discrimination is fomented when an applicant's race (or gender) is evident to a potential home seller (or employer, etc.). At least two cues to race and gender act as triggers: visual and auditory. These two stimuli influence a minority's success when entering into, and advancing within, the housing market. This article examines the linguistic nature of housing discrimination among minority groups, studying the nature of auditory discrimination of racial speech cues. We are particularly interested in the possibility of auditory discrimination in the absence of any visual cues and in determining the existence of micro-linguistic (i.e., phonetic) markers of dialect (Labov, 1972b). Our preliminary findings from four experiments indicate that (a) dialect-based discrimination takes place, (b) ethnic group affiliation is recoverable from speech, (c) very little speech is needed to discriminate between dialects, and (d) some phonetic correlates or markers of dialects are recoverable from a very small amount of speech.

Throughout this article we discuss three broad dialects of American English: African American Vernacular English (AAVE), Chicano English (ChE), and Standard American English (SAE). These dialects are chosen because of data availability and the fairly strong ethnic group affiliation tied to the dialects (confirmed in the second experiment described below). Our experimental results suggest the presence of parameters for phonetic articulation of each dialect learned alongside phonological contrasts. Because dialects consist, in part, of accents that constitute learned features of the phonological and phonetic systems of a language variety, holding phonology constant and comparing phonetic implementation across dialects provide a method to reveal the prominent phonetic features of each dialect.

We began by asking whether dialect discrimination is possible by using phonetic cues alone, and if it is possible, what cues trigger discrimination. Given that discrimination according to race or national origin is illegal under the Fair Housing Act and the Civil Rights Act of 1968 as amended, we seek to demonstrate that the identity of race (or national origin) is reflected, not only visually but also auditorily in an individual's speech. In addition, we endeavor to establish that listeners hear and positively identify a speaker's dialect with great accuracy. The following experiments reveal the possibility of auditory discrimination and the probability of social discrimination by auditory identification of dialects. Auditory cues thus comprise a significant factor in establishing evidence in effecting a prima facie case where the Fair Housing and Civil Rights Acts are involved.[1]

EXPERIMENT 1

Baugh's personal experiences while trying to rent an apartment in the San Francisco area provided an impetus and opportunity to study dialect identification and discrimination. This portion of our research addresses the issue of whether housing discrimination is exhibited in the absence of visual cues. Most legal cases appealing to the Fair Housing and the Civil Rights Acts try to establish that the defendant discriminated against the plaintiff because of obvious visual cues. In addition to this existing standard of proof, this experiment indicates that housing discrimination arises in the absence of these visual cues. In predominantly White geographic locales where discrimination against minorities would potentially be the greatest, the percentage of appointments secured to view housing is less than chance for callers using nonstandard dialects. Projecting this to the population at large, the evidence shows that a member of a minority group is much less likely to get an appointment to see an apartment in these White locales, even when he or she is qualified to purchase or rent in those areas. In these examples, auditory discrimination arises without visual contact.

Our research pertains to discrimination litigation by addressing issues that might be raised in the burden-of-proof tests that the plaintiff has to show (established in *McDonnell Douglas Corporation v. Green*, 1973). The plaintiff has to prove (a) "discrimination by a preponderance of the evidence," and (b) once the defendant has shown that he or she behaved toward the plaintiff in a nondiscriminatory fashion, that this behavior was simply pretext and false.

In building the evidence against the accused, a plaintiff establishes a prima facie case under the Fair Housing and Civil Rights Acts in several ways. The first item the plaintiff must prove is that he or she is a member of a racial minority. In addition, the plaintiff must verify that he or she applied for and was qualified to rent or purchase certain property of housing. The plaintiff must also demonstrate that he or she was rejected in some manner. The last proof the plaintiff must provide is that the housing or rental property remained unrented thereafter.

We are compelled to understand the possiblity of auditory discrimination from cases like *HUD v. Ross* (1994). In this suit, Judge Cregar notes the role accent played in the outcome of the case.

It is undisputed that Magaly Dejesus and Teresa Sanchez are Hispanic. Their distinct Hispanic accents clearly revealed their national origin to Mr. Ross. Although neither filled out a rental application, Mr. Ross did not afford them the opportunity to do so. He hung up on Ms. Dejesus when he learned that she received AFDC, and he never returned the message Ms. Sanchez left with Mr. Ross' secretary. (p. 8)

This example and a few others (e.g., *City of Chicago v. Matchmaker Real Estate Sales Center*, 1992) with respect to Ms. Frazier's "accent") demonstrate that individuals are capable of being held liable because of their auditory assessment of a speaker.

However, not all jurists agree that racial identity is ascertained in the absence of visual prompting. Judge Heifetz in his discussion of *HUD v. Cox* (1991) notes,[2]

> Even if Ms. Rancatti did not see Mr. Edwards, Ms. Oliver or her daughter, the Secretary conjectures that Ms. Rancatti knew Mr. Edwards' race from his "distinctive vocal characteristics." However, while his voice may have a distinctive timbre, upon hearing his testimony I was unable to discern, nor was there any other evidence upon which to conclude, that his voice or speech was characteristic of any particular racial or ethnic background. All that can be said about any distinction in his speech is that an occasional word revealed his Brooklyn heritage. (pp. 10-11)
>
> Even if [Mr. Edward's] voice revealed an ethnic or racial background, it is questionable whether such lineage could have been distinguished, given the intercom's poor quality. (p. 11, note 10)

For our first experiment, the null hypothesis is that there is no significant difference in appointments made by locale by dialect. The test hypothesis entails a relation between the racial and ethnic constituency of a geographic area and the success in establishing an appointment by dialect type. Census figures indicate that in the San Francisco metropolitan area, which ranks 44th in the percentage of African American households in the nation, the African American to White ratio in the value of homes is 0.64 (U.S. Census Bureau, 1990). This ratio reflects the national ratio of 0.61. African Americans have a median home value of $223,200, whereas Whites have a median home value of $348,200. These median home values are the highest of the Top 50 metropolitan areas. However, we show below that evaluation of segregation by major metropolitan areas is uninsightful. Instead, a local, city-by-city level examination of segregation is more profitable than a more general one.

METHOD

Baugh conducted the telephone interviews in person. Because Baugh (who is African American) grew up in inner-city communities in Philadelphia and Los Angeles, he is personally familiar with AAVE, ChE, and SAE dialects. This use of a tridialectal speaker controls for cross-speaker variation. In this respect, this study differs from other such guise studies as Lambert and Tucker (1972) and Tucker and Lambert (1975), which use different speakers in experiments determining attitudes toward dialects.

Prospective landlords in five distinct locales were identified by classified advertisements in regional newspapers. The number of calls for the geographic areas is unbalanced and is determined by housing availability over time for each locale. The geographic location of each property was noted. Baugh telephoned the landlords on three separate occasions, randomly using each dialect in different sequences with no less than 30 minutes between calls. Each call began with the phrase, "Hello, I'm calling about the apartment you have advertised in the paper." Different return telephone numbers were used for each dialect, along with different pseudonyms. This procedure of anonymity parallels legally approved practices of testers used by the Department of Housing and Urban Development and similar organizations when suspecting discriminating practices by landlords (e.g., *City of Chicago v. Matchmaker Real Estate Sales Center*, 1992; *Johnson v. Jerry Pals Real Estate*, 1973; *United States v. Youritan Construction Company*, 1975).

RESULTS AND DISCUSSION

The results show a clear pattern of potential discrimination associated with the three dialects by geographic area. Thus, we reject the null hypothesis and accept the experimental hypothesis. Tables 1, 2, and 3 show that the percentage of appointments made in each locale corresponds approximately with the ethnic makeup of the geographic area. Tables 2 and 3 display 1990 census data for percentage of population and percentage of householders who belong to particular racial and ethnic groups (U.S. Census Bureau, 1990). If the null hypothesis is rejected (i.e., if the positive appointment rate hovered around chance), then we expect a 50% success rate for each cell in Table 1. Percentages above and below the 50% mark indicate a variance from chance. By examining all three tables, we observe that in the traditionally White areas, Woodside and Palo Alto, the strongest bias is against the nonstandard dialects.

Distinct patterns appear when we compare the population of each locale with the number of householders with appointments secured by dialect type. Where SAE is concerned (see Figure 1), we see a relatively level pattern across the locales regardless of population and householder density (60% to 70%). For the nonstandard dialects—AAVE (see Figure 2) and ChE (see Figure 3)—minority population and householder information is paralleled by the percentage of housing appointments.

Compared to a segregation study using 1980 census data of the San Francisco area (Massey & Fong, 1990), the AAVE dialect pattern is understandable given the index of dissimilarity of African Americans to Whites (0.717, where 0 represents similarity and 1 represents dissimilarity). However, whereas Hispanics were found to be more

Table 1
Confirmed Appointments to View Apartments Advertised for
Rent in Different Greater San Francisco Geographic Areas (in percentages)

Dialect Guise	Geographic Area				
	East Palo Alto	Oakland	San Francisco	Palo Alto	Woodside
AAVE	79.3	72.0	63.5	48.3	28.7
ChE	61.9	58.3	53.2	31.9	21.8
SAE	57.6	68.7	71.9	63.1	70.1
Total number of calls for each locale	118	211	310	263	87

Note. AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English.

Table 2
Population in Different Greater San Francisco
Geographic Areas by Race and Ethnicity (in percentages)

Population	Geographic Area				
	East Palo Alto	Oakland	San Francisco	Palo Alto	Woodside
African American	42.9	43.9	10.9	2.9	0.3
Hispanic	36.4	13.9	13.9	5.0	3.8
White	31.7	32.5	53.6	84.9	94.7

Source. U.S. Census Bureau (1990).

Table 3
Householders in Different Greater San Francisco
Geographic Areas by Race and Ethnicity (in percentages)

Householder	Geographic Area				
	East Palo Alto	Oakland	San Francisco	Palo Alto	Woodside
African American	47.1	43.2	10.0	2.4	0.2
Hispanic	23.4	9.6	10.1	3.7	1.8
White	34.6	39.7	64.9	88.1	97.0

Source. U.S. Census Bureau (1990).

similar to Whites (0.402) than African Americans, in the present experiment, ChE guises not only follow the general trend of AAVE guises but also have the lowest percentage success rate in securing an appointment. This is even true in the geographic areas where the number of Hispanics exceeds the number of African Americans (San Francisco, Palo Alto, and Woodside).
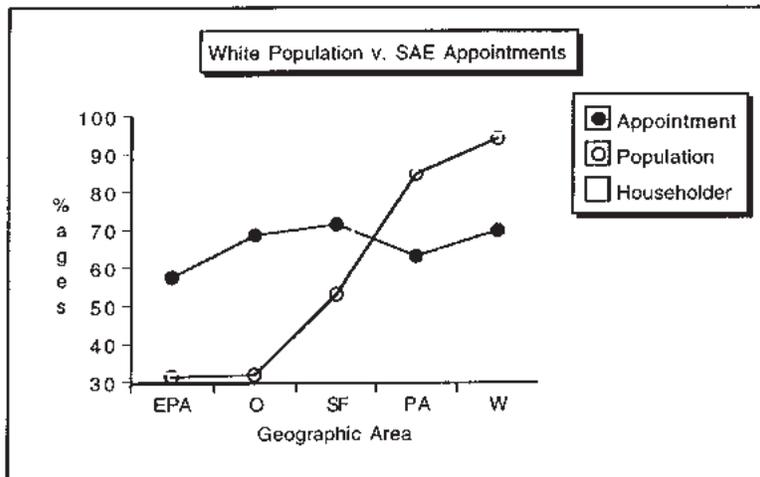
Figure 1. **Comparison of White population and householder data to percent-
age positive SAE appointments.**
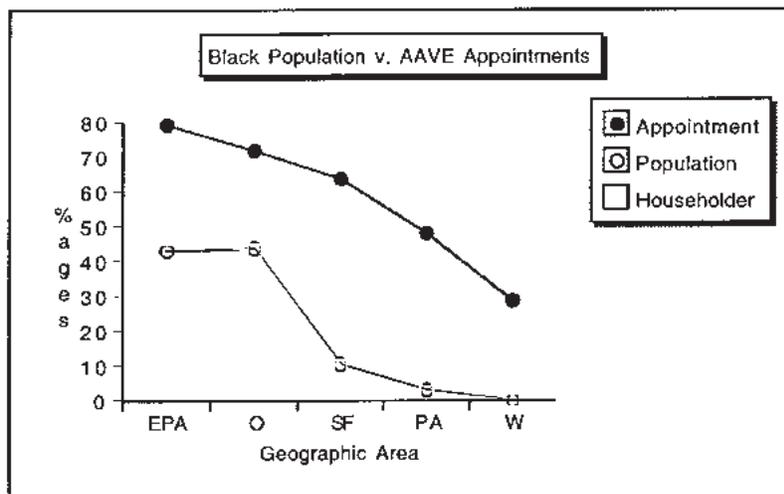*Note.* SAE = Standard American English.

Figure 2. **Comparison of African American population and householder data
to percentage positive AAVE appointments.**
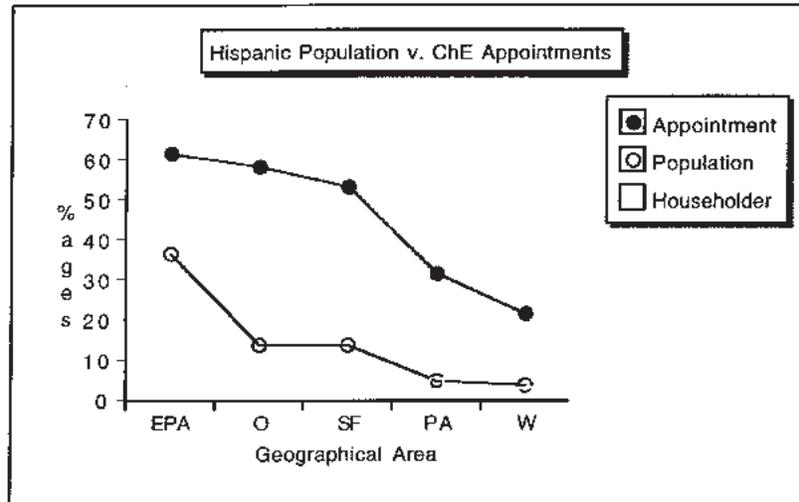*Note.* AAVE = African American Vernacular English.

Figure 3. **Comparison of Hispanic population and householder data to percentage positive ChE appointments.**
Note. ChE = Chicano English.

We conclude, given the evidence of this first experiment, that auditory cues constitute stimuli for disparate impact and nonaccidental disparate treatment cases. Disparate treatment discrimination cases involve "discriminatory motives" established through the discriminator's

> knowledge of the complainant's minority status, the very factor on which discriminatory animus relies. Disparate treatment is volitional, not accidental. It is an action of the will, having as its end a conscious choice among known alternatives. Therefore, knowledge of the alternatives is a necessary predicate for the exercise of that will. (*HUD v. Cox*, 1991, pp. 9-10)

Disparate impact cases involve indirect, and often unintentional, cases of discrimination. Furthermore, there may be reasons for the paucity of civil rights cases in which the burden of proof rests at least indirectly on auditory cues, as compared to the preponderance of cases that involve visual ones. One possibility is that the auditory kinds of discrimination are difficult to monitor. In addition, potential householders may scarcely suspect that they are being discriminated against; the landlord subtly discriminates by informing the minority speaker that there are no vacant apartments. This is supported by a survey reported in Clark (1993), indicating that the primary source of discrimination is

the individual landlord. The upshot of variation by neighborhood within a general metropolitan area is that it corroborates evidence that segregation is fairly local (Massey & Hajnal, 1995).

EXPERIMENT 2

Given that Baugh is tridialectal and might favor one dialect over the others as a default dialect, we might wonder whether the guises in the first experiment are representative of the appropriate racial group and not exaggerated stereotypes instead. Following the first experiment, a series of experiments on "ethnic identification" evaluations were conducted. The next experiment is an attempt to understand whether dialect identification is possible at the macro-linguistic or sentential level. In contrast, subsequent experiments explore identification at the micro-linguistic level.

Much is already known about the sentence- and morpheme-level differences between SAE and AAVE, along with the phonemic and lexical alternations (see, e.g., Baugh, 1983; Dillard, 1972; Labov, 1969, 1972a; Wolfram, 1969). Several of these phonological and morphophonemic differences could possibly affect our research. One of the notable characteristics of AAVE that is different from SAE is the absence in AAVE of certain sonorants, /r l n/, in syllable coda position. Other sounds, as well, may be absent from AAVE, for example, /-s/ suffixes (plural, third person singular, possessive) and consonants (particularly /t d/) from consonant clusters. AAVE also exhibits final obstruent devoicing and consonant mergers, such as [θ ~ f].

Another important nonstandard dialect in the United States is Chicano English (ChE), which is also part of our investigation. Several studies examine the linguistic differences between ChE and SAE (e.g., González, 1988; Penfield, 1984; Penfield & Orstein-Galicia, 1985; Wald, 1984). Particularly important is the phonetic study of Godinez (1984). Again, we need to pay close attention to the phonological and morphophonemic differences between the dialects. Intonation on utterances of all sizes differs between ChE and SAE. For example, utterances in ChE begin at a higher pitch, although ChE intonations pattern more closely with English than with Spanish. Segmental changes differentiate SAE and ChE, for example, palatal interchange, fricative and affricate devoicing, and labiodental fricatives merging into coronal stops. Like AAVE, ChE modifies certain consonant clusters, especially initial and final clusters involving /s/. Unexpectantly, however, the mean duration of ChE vowels is more closely aligned with SAE vowels than with Spanish vowels.

METHOD

Stimulus tokens for this experiment were recorded by speakers of the three target dialects. The number of speakers, totaling 20, varied across the racial and ethnic groups. In addition, Baugh recorded tokens in each of the three dialects. Each token consisted of the sentence, "Hello, I'm calling to see about the apartment you have advertised in the paper." The tokens were then randomized for presentation.

In the experimental stage, 421 undergraduate and graduate students at Stanford (382 native speakers of English, 39 nonnative speakers) listened to each token once without response. The students then listened to the tokens a second and a third time, indicating two presumed traits in a forced-choice experiment. The two traits students were asked to evaluate the listeners for were the race/ethnicity and gender of speakers. The possible answers for race and ethnicity were "African American," "Hispanic American," and "European American." Combined with the two choices of gender, participants selected one of six possible responses for each token.

Given that each token has the possibility of being assigned one of six choices, the null hypothesis is that each guise should be identified correctly 16.6% of the time. We predict instead that tridialectalism is likely and that these guises are identifiable at the same rate as nontridialectal ones.

RESULTS AND DISCUSSION

The results of this study indicate that participants systematically identified Baugh's guises as being produced by an African American male (i.e., using AAVE), a Latino (i.e., using ChE), or a White male (i.e., using SAE) (see Table 4). Thus, we reject the null hypothesis and accept the test hypothesis.

Guise identification, as expected, is possible at the macro-linguistic level. All three guises are judged as being representative of the target dialect (slightly more than four out of five times for the least-identified guise). These macro-linguistic cues to dialect present an advantage as they are overt indicators of a speaker's ethnic identity. The problem, though, that this study faces is in explaining micro-linguistic, or more subtle, cues.

EXPERIMENT 3

To determine the feasibility of investigating the phonetics of dialects, a second perceptual experiment was conducted in which we tested listeners' ability to recognize dialects at the micro-linguistic, or

Table 4
Dialect and Racial Identification

Dominant Dialect/Racial Identification	Guise or Gender	% Correct Identification
AAVE/African American	Male	97
	Male	95
	Female	85
	Baugh (AAVE)	84
	Male	77
ChE/Hispanic American	Baugh (ChE)	91
	Male	86
	Female	79
SAE/European American	Male	92
	Female	87
	Baugh (SAE)	86
	Female	86
	Female	83
	Male	81

Note. AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English.

phonetic, level. The results are intended to contribute to our under-standing of the psychological processes enabling phonetic distinctions.

This third experiment deals with two basic issues of dialect production and perception. Dialects differ in their syntactic, morphological, or semantic subcomponents, but our concern is to understand the role played by phonetics and phonology. First, we ask, "How do dialects differ in pronunciation?" and second, "How do listeners identify dialects by pronunciation?" Because discrimination may crop up in telephone conversations, this experiment helps us better understand the cognitive role phonetics plays in establishing listeners' beliefs about a speaker's racial identity.

Phonetic features in the speech stream not distinguishing words (the "noncontrastive" features) are used for speaker and dialect recognition. Whether a given feature is contrastive or noncontrastive is a language-particular (and dialect-particular) choice and therefore must be learned. The learning results in shared knowledge about society in general, interlocutors' positions in society, and appropriate discourse norms given the discourse situation (Baugh, 1983). Therefore, the particular acoustic features signaling a dialect distinction cannot be predicted by racial genetics. In other words, the dialect features are learned speech characteristics, rather than being anatomically determined. One consequence of having to learn the phonetic grammar of a dialect is that individual speakers can control several dialects. Competency in more than one dialect is quite common, especially among speakers of nonstandard varieties. Thus, among speakers of AAVE, the presence or absence of /-s/ suffixes is tied to familiarity and dialect

group membership, rather than to racial characteristics of the speaker (Baugh, 1983).

Breathiness is one example of a phonetic feature that serves a contrastive grammatical function or, when noncontrastive, behaves like a sociolinguistic marker. Breathiness is used contrastively in such languages as Hindi and Marathi to distinguish different words but is not used contrastively in SAE (Ladefoged & Maddieson, 1996). Klatt and Klatt (1990) found, however, that female SAE speakers are on average more breathy than male speakers. Therefore, breathiness has different status in different languages. That is, breathiness has potentially different cognitive functions across languages. The breathy voice of female speakers of SAE is a learned social behavior and is a marker of gender in SAE.

The acoustic signal carries a variety of information about the individual speaker beyond just the phonemic content of the signal. Some of the acoustic features of the signal are produced by the gross anatomical characteristics of the speaker's vocal tract. For example, differences in physiology, such as vocal tract length and vocal fold density, influence the average fundamental frequency of men and women. The effects of such gender differences entail higher fundamental and formant frequencies for women as compared with men (Hollien, 1962; Klatt & Klatt, 1990; Peterson & Barney, 1952). These physiological features are not controlled by the speaker and therefore do not represent different learned dialects. Regarding the physiology of African Americans and Whites, Sapienza (1997)—in contrast to Boshoff (1945, cited in Walton & Orlikoff, 1994)—found that there is no significant difference between the groups for many laryngeal aerodynamic and acoustic characteristics.

The null hypothesis in this experiment is that there is no difference between the dialects by identification. That is, each dialect should display recognition at the level of chance. Instead, we predict that the phonetic characteristics in a short portion of speech are sufficient to trigger identification across the dialects.

METHOD

For this experiment, only the word *hello* from Baugh's single-sentence utterances spoken in AAVE, ChE, and SAE were used. The word was extracted from the sentence, "Hello, I'm calling about the apartment you have advertised in the paper." We have several reasons for examining one word. This allowed us to hold external factors to a minimum. Second, it also illustrates how little speech is needed for dialect identification. "Hello" is a self-contained utterance, making perceptual studies more natural. By focusing on one short word, we are able to hold utterance duration well below one second ($\bar{x} = 414$ msec.), making it comparable to other studies (e.g., Walton & Orlikoff, 1994).

The word *hello* neutralizes lexical, syntactic, and phonological differences across dialects. In other words, it lacks the environment in which we expect other dialectal variations.

This experiment was conducted during two semesters, Spring and Fall 1997. For this study, we used 50 undergraduates at the University of Delaware (Spring 1997: 30; Fall 1997: 20). All of the participants were Caucasian native speakers of SAE. Ten instances of "hello" repeated twice for each of the three dialects comprised one block of data. These 60 tokens were randomized. Each participant was twice presented with the block of data so that a total of 120 tokens were presented to each participant. During a 2-second pause, participants indicated, for each token, which dialect they believed they heard. The data below are combined from the two iterations of the experiment.

RESULTS AND DISCUSSION

Statistical procedures reject the null hypothesis in this experiment. First, the Accuracy Index (AI) reflects the overall pattern where more responses lie on the diagonal of a 3×3 matrix (represented by the bolded cells in Table 5). In the following confusion matrices, the AI is the sum of the diagonal cells, divided by the sum of all cells. Thus, the higher the AI value is, the better the responses reflect an ability to identify tokens among the dialect types, whereas the lower the AI value is, the responses become closer to chance. Second, we reject the null hypothesis by examining each of the individual cells relative to the total number of tokens. For reference, Table 5 shows that when identification is by chance, each stimulus category receives an equal number of responses (11%); when identification is "perfect," cells a, e, and i receive the maximum number for that stimulus (33%).

The results show that, overall, participants are able to successfully identify tokens among the dialects when only hearing the word *hello*. We reject the null hypothesis in favor of the experimental hypothesis because in Table 6 the overall AI is .72, significantly better than chance. In other words, respondents correctly identified between the dialects more than 70% of the time. The individual cells do not reflect a response by chance either. Diagonal cells display response rates of 15%, 27%, and 29%, across the three dialects. The cells displaying a response closest to chance are the AAVE tokens identified as AAVE (15%) and as SAE (14%).

A comparison of the misidentification cells (cells b, c, d, f, g, h in Table 6) shows two things. (The pattern of misidentification is shown in Table 7.) First, when AAVE tokens are misidentified, AAVE tokens tend to be misidentified as SAE tokens. Second, ChE and SAE tokens are misidentified as AAVE tokens. In general, however, the rates of misidentification differ, with AAVE exhibiting the highest misidentification of the three dialects.

Table 5
General Confusion Matrices

Response	Stimuli AAVE	ChE	SAE	"Chance"			"Perfect"		
AAVE	**a**	b	c	11	11	11	33	0	0
ChE	d	**e**	f	11	11	11	0	33	0
SAE	g	h	**i**	11	11	11	0	0	33

Note. AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English.

Table 6
Confusion Matrix and Summary Statistics by Dialect

Response	Stimuli AAVE			ChE			SAE			Row Total	
AAVE	a	923	(15%)	b	280	(5%)	c	196	(3%)	1,399	(23%)
ChE	d	235	(4%)	e	1,607	(27%)	f	41	(1%)	1,883	(31%)
SAE	g	842	(14%)	h	113	(2%)	i	1,763	(29%)	2,718	(45%)

Note. AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English. $\chi^2 = 4{,}510$, $df = 4$, $p < .001$; Accuracy Index (AI) = .72; percentages = percentage of total for that cell.

More than one factor, then, must account for these misidentification patterns. Factors might include (a) a response bias toward the SAE dialect as a neutral, default dialect; (b) a standard versus nonstandard bias favoring the AAVE-ChE misidentification; or (c) a phonetic similarity favoring the AAVE-SAE misidentification.

If we take a response bias into account, the expected values necessary for computing the chi-square value are adjusted (see Table 8). In standard statistical packages, the response bias correction is calculated automatically (adjusted by "marginal values," i.e., the sum of each column and row). With these adjusted expected values, only the AAVE tokens misidentified as SAE tokens (14% in cell g) come close to chance. The response rate for AAVE tokens perceived correctly as AAVE tokens is almost twice chance (15% v. 8%). To account for this, we applied a response bias adjustment.

Furthermore, the standard dialect contrasts with the nonstandard dialects. The comparison shown in Table 9 highlights a response bias by grouping AAVE and ChE, and comparing the nonstandard group with the more standard SAE dialect.

At this point, we have established that listeners are capable of discriminating among dialects and that this discrimination is eased by a low-level identification of the dialects in a short amount of time. What

Table 7
Stimulus and Response Misidentifications

Response	Stimuli		
	AAVE	ChE	SAE
AAVE	**a**	b	c
ChE	d	**e**	f
SAE	g	h	**i**

Note. AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English.

Table 8
Row-Adjusted Confusion Matrix

Response	Stimuli		
	AAVE	ChE	SAE
AAVE	8	8	8
ChE	10	10	10
SAE	15	15	15

Note. AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English.

Table 9
Confusion Matrix and Summary Statistics for Standard and Nonstandard Dialects

Response	Stimuli			
	Nonstandard		Standard	
Nonstandard	3,045	(51 / 36%)	237	(4 / 18%)
Standard	955	(16 / 30%)	1,763	(29 / 15%)

Note. $\chi^2 = 2,223$, $df = 1$, $p < .001$; Accuracy Index (AI) = .80; percentages = percentage of total responses and the expected percentage including the adjustment for response bias (see text).

still remains is at least a partial explanation of what phonetic features of the speech stream act as sociolinguistic markers.

EXPERIMENT 4

We performed a variety of acoustic measurements on the same "hello" data used in the perceptual experiment (Experiment 3). In this acoustic experiment we looked for acoustic differences between the dialects to determine cues listeners use to identify dialects. Our

answer, from the acoustic measurements, is that at least four acoustic cues are viable for distinguishing at least one dialect from the other two: the frequency of the second formant in the /ɛ/, the location in the word where the pitch reaches a peak, the duration of the first syllable /hɛ/, and the harmonic-to-noise ratio (HNR).

In this acoustic experiment, the null hypothesis is that there is no difference between the dialects for any of the phonetic measurements we perform. We predict, instead, that there is a significant distinction among the dialects accounting for why the tokens are recognized so well.

METHOD

Following other studies (e.g., Klatt & Klatt, 1990; Stevens & Hanson, 1995; Walton & Orlikoff, 1994), we measured each instance of "hello" for several acoustic characteristics. We measured the segment, syllable, and word durations (and ratio of these durations to the duration of the word). Next, we measured the midpoint formant frequencies (F1, F2) for each vowel. The amplitudes of the first two harmonics for each vowel (|H1|, |H2|) and the ratio of |H1| to |H2| were also measured and computed. We then measured the midpoint pitch of each vowel and the location in the word of the highest F0 peak. A ratio of the duration from the beginning of the word to the location of the pitch peak was computed. HNR was the final calculation for each word. The 30 tokens (10 for each dialect) were compared on 28 variables by running a three-way analysis of variance (ANOVA) for the different variables.

RESULTS AND DISCUSSION

Four measurements are significant in distinguishing the three dialects: the frequency of the second formant in /ɛ/, the pitch peak ratio, the duration of the first syllable, and HNR. However, none of these features reliably differentiate all three dialects. Table 10 shows the matching results of Duncan and Tukey post hoc test. Only the value for F2 in /ɛ/ is found to be significantly different in a Scheffé test for distinguishing ChE and AAVE from SAE. Therefore, the value of F2 in /ɛ/ is the best cue we found. Nevertheless, the analysis remains incomplete because no factor accounts for the misidentification of AAVE tokens as SAE (the shift from cell a to g in Table 7) or SAE tokens as AAVE (the shift from cell i to c in Table 7).

First, as seen in Table 11 we reject the null hypothesis for the frequency of the second formant in /ɛ/ (SAE = 1,195 Hz, ChE = 1,498 Hz, AAVE = 1,445 Hz). We interpret this characteristic as an indication that the two nonstandard dialects tend to raise and front the front vowel /ɛ/ towards /e/. Again, it is worth pointing out that this characteristic is the strongest of the measurements because it is the only one in

Table 10
Acoustic Measures Differentiating Dialects

	F2 in /ɛ/	F0 Peak Ratio	/hɛ/ Duration	HNR	Misidentification	Cell
ChE versus SAE	*				Least misidentified	(cells f, h)
AAVE versus ChE		*	*	*		(cells b, d)
AAVE versus SAE	*		*	*	Most misidentified	(cells c, g)

Note. HNR = harmonic-to-noise ratio; AAVE = African American Vernacular English; ChE = Chicano English; SAE = Standard American English.
*$p < .05$.

which the significance pattern is the same in Duncan, Tukey's honestly significant difference (HSD), and Scheffé post hoc tests. In addition, this measure of F2 involves a fairly strong relation as indicated by a ω^2 value of .369. That is, 36.9% of the variance across the dialects is explained by the frequency of the second formant alone.

This factor is not without problems though. Because SAE is significantly different from ChE and AAVE, this factor might explain why ChE tokens are misidentified as AAVE tokens (the shift from cell e to b in Table 7). However, it does not explain the failure of AAVE tokens to be misidentified as ChE tokens.

Second, the relative location of the pitch peak in the word proved significantly different across the dialects (ChE = 56%, SAE = 53%, AAVE = 49%). The ANOVA results are shown in Table 12. This characteristic reflects an alternate placement of stress from the final syllable (ChE, SAE) to the initial syllable (AAVE). In other words, ChE tokens reflect a final stress ([heló]), whereas the AAVE tokens reflect a more initial stress ([hέlo]). The ω^2 value is .135; that is, 13.5% of the variance between dialects is accounted for by the location of the pitch peak relative to the word alone.

The location of the pitch peak as a defining characteristic across the dialects is uncertain as well. First, it failed to display significance in the Sheffé test. Second, it fails to account for token misidentifications. For instance, if ChE and AAVE are significantly different from each other in the location of the pitch, then the observed pattern of ChE tokens being misidentified as AAVE should be dispreferred by participants.

Third, Table 13 shows that the duration of the first syllable /hɛ/ (AAVE = 115 msec, SAE = 90 msec, ChE = 88 msec) is also significant across dialects. The ω^2 value is .16; that is, 16.0% of the variance between dialects is accounted for by the duration of the first syllable alone.

This characteristic has problems, especially because it is an absolute value measurement and not a ratio in reference to the duration of the word. Using the same logic as in our discussion of the problems

Table 11
ANOVA for F2 in /ɛ/

Source	df	SS	MSS	F-Value
Between groups	2	525,391.27	262,695.63	9.79*
Within groups	27	724,796.60	26,844.32	
Total	29	1,250,187.87		

Note. ANOVA = analysis of variance; *SS* = sum of squares; *MSS* = mean sum of squares.
*p < .05.

Table 12
ANOVA for Pitch Peak

Source	df	SS	MSS	F-Value
Between groups	2	0.0258953	0.0129476	3.34*
Within groups	27	0.1045779	0.0038733	
Total	29	0.1304732		

Note. ANOVA = analysis of variance; *SS* = sum of squares; *MSS* = mean sum of squares.
*p < .05.

Table 13
ANOVA for Duration of First Syllable, /hɛ/

Source	df	SS	MSS	F-Value
Between groups	2	4,691.1183	2,345.5591	3.86*
Within groups	27	16,415.8310	607.9937	
Total	29	21,106.9492		

Note. ANOVA = analysis of variance; *SS* = sum of squares; *MSS* = mean sum of squares.
*p < .05.

with the value for F2 of /ɛ/ above, we might be led to believe that SAE and ChE would be misidentified. On the contrary, these two dialects are misidentified the least. That is, in Table 7, we do not observe cells e and i being misidentified in cells h and f. This is an odd finding because the percentage duration of the first syllable relative to the duration of the word is not significant.

Fourth, HNR (AAVE = 12.1 dB, ChE = 13.6 dB, SAE = 13.9 dB) is the ratio of noise relative to the harmonic structure of a wave. It, too, is significantly different across the dialects (see Table 14). The ω^2 value for HNR is .145, that is, 14.5% of the variance between dialects is accounted for by the amplitude of noise in the wave relative to the amplitude of the harmonics.

Table 14
ANOVA for Harmonic-to-Noise Ratio (HNR)

Source	df	SS	MSS	F-value
Between groups	2	17.742325	8.871163	3.55*
Within groups	27	67.485995	2.499481	
Total	29	85.228320		

Note. ANOVA = analysis of variance; SS = sum of squares; MSS = mean sum of squares.
*$p < .05$.

HNR, like the other three features, is difficult to interpret. This feature has many possible articulatory correlates. HNR simply measures the amount of noise, and many factors lead to increased noise. Therefore, HNR does not provide enough information to allow reasonable articulatory interpretations. HNR is used clinically to diagnose a sizable number of laryngeal dysfunctions. These dysfunctions lead to significant morphological differences between normal and abnormal larynges. In our data, HNR does not vary drastically, but it is significantly different for SAE and AAVE voices. Furthermore, the HNR results provide the same dialect separation as the first syllable duration results. Given the choice between the readily interpretable duration difference and the much vaguer HNR results, we view the first syllable duration as a more satisfactory sociolinguistic marker of dialect.

CONCLUSION

The experiments described in this article link housing opportunities with dialect use. Housing discrimination induced by speech characteristics does take place. Dialects are discriminated by normal listeners. Very little speech is required for dialect identification—a single word suffices. Dialects are discriminated with acoustic phonetic measures. Patterns of perceptual misidentification point to a multiplicity of factors for further study.

Looking back at the experiments as a whole, we should wonder whether the acoustic characteristics of the guises influenced the outcome of the discrimination survey. Consider that in Table 1 percentages of appointments made using ChE guises are the lowest in four of five geographic areas. Although we might be led to believe from this result that Hispanic Americans experience more discrimination than African Americans, we could well be misled. In Table 6, however, ChE tokens are identified much better than AAVE tokens. Putting the two observations together, we see a possibility that ChE tokens as produced by Baugh are more salient as exemplars of nonstandard speech than his AAVE tokens, and they are thus less likely to be confused with the SAE tokens.

This research will be continued with expanded pools of speakers and listeners, including tri-, bi-, and monodialectal speakers. More acoustic measures will be examined, concentrating on acoustic measures that have clear articulatory sources. Research of the type presented here assesses trade-offs in the acoustic cues for dialect identification by using synthesized stimuli.

NOTES

1. What we are suggesting goes beyond what is covered under rules of evidence (Graham, 1987). Rule 901(b)(5) states that voice identification is permissible as evidence to identify the voice. We propose that rules of evidence govern the identification of race by way of a speaker's voice. Thus, if a plaintiff shows that she has a voice representative of a nonstandard dialect and that the defendant can ascertain when any voice possesses characteristics of the dialect under review, then evidence is established in favor of the plaintiff.

2. In fairness to the parties involved, the judge ruled in favor of the defendants because of a preponderance of evidence indicating that Ms. Rancatti had neither acted in a racist fashion nor acted habitually in such manner.

REFERENCES

Baugh, J. (1983). *Black street speech: Its history, structure and survival*. Austin: University of Texas Press.

Boshoff, P. (1945). The anatomy of the South African Negro larynx. *South African Journal of Medical Sciences*, *10*, 113-119.

City of Chicago v. Matchmaker Real Estate Sales Center, 982 F.2d 1086 (7th Cir. 1992).

Civil Rights Act, § 801 *et seq*. (1968)

Clark, W. (1986). Residential segregation in American cities: A review and interpretation. *Population Research and Policy Review*, *5*, 95-127.

Clark, W. (1993). Measuring racial discrimination in the housing market: Direct and indirect evidence. *Urban Affairs Quarterly*, *28*, 641-649.

Courant, P. (1978). Racial prejudice in a search model of the urban housing market. *Journal of Urban Economics*, *5*, 329-345.

Dillard, J. (1972). *Black English: Its history and usage in the United States*. New York: Random House.

Fair Housing Act, 42 U.S.C. § 3601 *et seq*. (1968).

Galster, G., & Keeney, W. (1988). Race, residence, discrimination, and economic opportunity: Modeling the nexus of urban racial phenomena. *Urban Affairs Quarterly*, *24*, 87-117.

Godinez, M. (1984). Chicano English phonology: Norms vs. interference. In J. Orstein-Galicia (Ed.), *Form and function in Chicano English* (pp. 42-48). Rowley, MA: Newbury House.

González, G. (1988). Chicano English. In D. Bixler-Márquez and J. Orstein-Galicia (Eds.), *Chicano speech in the bilingual classroom* (pp. 71-82). New York: Peter Lang.

Graham, M. (1987). *Federal rules of evidence in a nutshell*. St. Paul, MN: West.

Hollien, H. (1962). Vocal fold thickness and fundamental frequency of phonation. *Journal of Speech and Hearing Research*, *5*, 237-243.

HUD v. Cox (HUDALJ 009-89-1641-1, September 26, 1991) [On-line]. Available: http://www.hud.gov/alj/pdf/cox.pdf

HUD v. Ross (HUDALJ 01-92-0466-8, July 7, 1994) [On-line]. Available: http://www.hud.gov/alj/pdf/ross.pdf

Johnson v. Jerry Pals Real Estate, 485 F.2d 528 (7th Cir. 1973).

Klatt, D., & Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *Journal of the Acoustic Society of America*, *87*, 820-857.

Labov, W. (1969). Contraction, deletion, and inherent variability of the English copula. *Language*, *45*, 715-762.

Labov, W. (1972a). *Language in the inner city: Studies in the Black English vernacular*. Philadelphia: University of Pennsylvania Press.

Labov, W. (1972b). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Cambridge, MA: Blackwell.

Lambert, W., & Tucker, G. (1972). *Bilingual education of children: The St. Lambert experiment*. Rowley, MA: Newbury House.

Massey, D., & Fong, E. (1990). Segregation and neighborhood quality: Blacks, Hispanics, and Asians in the San Francisco metropolitan area. *Social Forces*, *69*, 15-32.

Massey, D., & Hajnal, Z. (1995). The changing geographic structure of Black-White segregation in the United States. *Social Science Quarterly*, *76*, 527-542.

McDonnell Douglas Corporation v. Green, 411 U.S. 792, 93 S.Ct. 1817, 36 L.Ed.2d 668 (1973).

Myers, S., & Chan, T. (1995). Racial discrimination in housing markets: Accounting for credit risk. *Social Science Quarterly*, *76*, 543-561.

Penfield, J. (1984). Prosodic patterns: Some hypotheses and findings from fieldwork. In J. Orstein-Galicia (Ed.), *Form and function in Chicano English* (pp. 49-59). Rowley, MA: Newbury House.

Penfield, J., & Orstein-Galicia, J. (1985). *Chicano English*. Philadelphia: John Benjamins.

Peterson, G., & Barney, H. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, *24*, 175-184.

Sapienza, C. (1997). Aerodynamic and acoustic characteristics of the adult African American voice. *Journal of Voice*, *11*, 410-416.

Stevens, K., & Hanson, H. (1995). Classification of glottal vibration from acoustic measurements. In O. Fujimura & M. Hirano (Eds.), *Vocal fold physiology: Voice quality control* (pp. 147-170). San Diego, CA: Singular Publishing Group.

Tucker, G., & Lambert, W. (1975). White and Negro listeners' reactions to various American-English dialects. In J. Dillard (Ed.), *Perspectives on Black English* (pp. 369-377). The Hague, the Netherlands: Mouton.

U.S. Census Bureau. (1990). Census of population and housing [On-line]. Available: http://govinfo.kerr.orst.edu/stateis.html

U.S. Census Bureau. (1995). Housing in metropolitan areas—Black households [On-line]. Available: http://www.census.gov/apsd/www/statbrief/sb95_5.pdf

United States v. Youritan Construction Company, 509 F.2d 623 (9th Cir. 1975).

Wald, B. (1984). The status of Chicano English as a dialect of American English. In J. Orstein-Galicia (Ed.), *Form and function in Chicano English* (pp. 14-31). Rowley, MA: Newbury House.

Walton, J., & Orlikoff, R. (1994). Speaker race identification from acoustic cues in the vocal signal. *Journal of Speech and Hearing Research*, *37*, 738-745.

Wolfram, W. (1969). *A sociolinguistic description of Detroit Negro speech*. Washington, DC: Center for Applied Linguistics.

Yinger, J. (1986). Measuring racial discrimination with fair housing audits: Caught in the act. *American Economic Review*, *76*, 881-893.