SOCIOPHONETIC APPLICATIONS OF SPEECH PERCEPTION EXPERIMENTS

ERIK R. THOMAS
North Carolina State University

ALTHOUGH STUDIES OF PERCEPTION are still largely assigned to the realms of experimental phonetics or psychology, sociolinguists have been recognizing the importance of perception. Several lines of experimental inquiry have emerged. Nevertheless, perception has been studied far less by sociolinguists than has speech production. One reason is that speech perception is daunting at first. Examining it requires careful attention to experimental design, a considerable amount of preparation, and, in many cases, use of a speech synthesizer. Even so, research on perception can be highly productive. This paper attempts to review the sorts of experiments that have been conducted in the past and to provide guidelines for sociolinguists interested in studying perception, with suggestions for future work.

Although perception has been a neglected stepsister of production in sociolinguistics, it, like Cinderella, may have its day soon. Two important factors could—and should—move perception to the forefront of sociophonetic research. One is simply the huge potential for sociolinguistic perception studies because the area has been neglected for so long. The other reason is a more practical one: although perception experiments require extreme attention to detail in the preparation phase, data analysis is generally less time-consuming than in production studies, and this difference may make it more attractive to researchers.

The aversion of much of sociolinguistics to perception has been, to some extent, more apparent than real. Many sociolinguistic studies over the past generation, especially instrumental studies, have succeeded in divorcing speech production from speech perception. However, perception issues may play a hidden role in studies that ostensibly address production. The reason is that variationists have not always carefully distinguished production from perception. This tendency is an artifact of the reliance of sociolinguistics on impressionistic transcription. The impressionistic tradition, based on the development of the International Phonetic Alphabet and of the Cardinal Vowel system of Daniel Jones, dominated dialect

American Speech, Vol. 77, No. 2, Summer 2002

geography, which was largely conceived before modern acoustic equipment was developed. This reliance has continued, for the most part, in sociolinguistics. The shortcoming of this approach is that, although linguistic atlas transcribers and sociolinguists believe that they are recording subjects' production, what they actually record is filtered through their own perceptual labeling abilities, practices, and strategies (Kerswill and Wright 1990).

An example of a hazard inherent in such a mixture involves one of the most frequently studied phenomena of English: the deletion of final stops in consonant clusters, as in *lift* pronounced [lɪf] or *desk* pronounced [dɛs]. Studies of consonant cluster simplification have customarily involved impressionistic transcription. However, Browman and Goldstein (1990) showed experimentally that, because of the overlapping nature of articulatory gestures, listeners may be unable to hear a stop in a consonant cluster, such as the [t] in *perfect memory*, even when an articulatory gesture is demonstrably present. Surprenant and Goldstein (1998) furthermore found predictable patterns in the perceptual masking of stops in clusters. These patterns matched many of the results reported in sociolinguistic studies of deletion of consonants in clusters (e.g., Wolfram 1969; Labov 1972; Guy 1980). The implication is that many of the stops tabulated in sociolinguistic studies as "deleted" were, in fact, articulated but were inaudible to the transcribers. Kerswill and Wright (1990) found that the opposite can happen as well: listeners may record an alveolar stop when none exists. Studies of consonant cluster simplification primarily describe the perception of transcribers and only secondarily the production of speakers. The binary interpretation of final stops as either present or deleted may describe perception satisfactorily, but it cannot truly describe production (see Janson 1983, 22–23, for another example).

Marriages of production and perception are certainly not invalid per se; as Lindblom (1980) notes, acoustic measurements of speech are meaningless unless they can be related to perceivable factors. However, variationists need to acknowledge more consistently the fact that production and perception do not always correspond neatly. Directing more studies at perception would help them to do so.

PREVIOUS SOCIOPERCEPTUAL APPROACHES

Sociolinguists and other investigators have utilized a variety of types of perception experiments to examine several sorts of sociolinguistic questions. Most work on perception in language variation has used spoken

words or phrases that were played to listeners for the purpose of testing one of five issues: (1) the ability of listeners to identify the regional dialect, ethnicity, or socioeconomic level of speakers; (2) how stereotypes can influence the perception of sounds; (3) the presence of vowel mergers or splits in perception; (4) how dialectal differences affect the categorization of phones; and (5) stereotypical attitudes, which are investigated by having subjects assess the personality of a speaker, the speaker's suitability for particular jobs, or other personal traits of the speaker. These issues are reviewed in the following sections, with the first split into three sections and the other four issues each represented by one section. This list of issues is somewhat limited, however, and many other objects of perceptual investigation are certainly possible. If the topic of socioperceptual studies is broadened to include not only synchronic language variation but also inquiry into the perceptual causes of diachronic change, especially by comparisons with misperceptions in a laboratory setting, additional studies could be added: see in particular John J. Ohala's work on the role of misperception in sound change (e.g., Hombert, Ohala, and Ewan 1979; Ohala 1981, 1985, 1989, 1993), but see also Browman and Goldstein (1991), who place more emphasis on the phasing of articulatory gestures, Jonasson (1971), and Foulkes (1997). Another important related topic to which perception experiments have been applied is the degree of accentedness of second-language or ethnic speakers (see, e.g., Brennan, Ryan, and Dawson 1975; Brennan and Brennan 1981; Sebastian and Ryan 1985). Further broadening of the topic could encompass studies correlating acoustic parameters of speech with emotional states (see, e.g., Goldbeck, Standke, and Scherer 1988).

IDENTIFICATION OF THE REGIONAL DIALECT OF SPEAKERS. Within this category, I include studies in which recordings of speakers of different dialects are played to subjects. Not included is "perceptual dialectology" or "folk dialectology," in which subjects demarcate dialect boundaries (usually on a map) or label dialects according to how "correct," "pleasant," and the like the subject considers them. Most studies of that type do not involve recordings. Examples of papers on perceptual dialectology include much of Dennis R. Preston's recent work (e.g., 1986, 1993a, 1993b, 1996) and the collection of articles in Preston (1999).

One of the first experiments involving dialect identification of recordings was Bush's (1967) study in which listeners were asked to identify the dialect of voices speaking American, British, or Indian English. The stimuli, which included nonsense words, real words, and sentences, were played either in original form, low-pass filtered, high-pass filtered, or center-

clipped. The filtering provided some indication about whether listeners could base their identifications on prosodic factors. As it turned out, listeners were generally able to identify the dialect of the filtered stimuli 75% or more of the time.

A more recent experiment is described in Preston (1993a, 359–67; 1993b, 41–46; 1996, 320–28). In this study, tapes of nine speakers who lived along a transect from Michigan to Alabama were played to two groups of subjects, one from southeastern Michigan and the other from southern Indiana. The subjects were asked to match the voices they heard with points along the transect, as shown on a map. They were fairly accurate in their assessments. However, the Michigan and Indiana subjects differed in which speakers they were best able to discriminate. Similar experiments have been conducted elsewhere. Stephan (1997) played recordings of 12 speakers of English dialects from around the world to German university students and found that American English was identified correctly most often, while South African and Welsh English were identified correctly least often. Williams, Garrett, and Coupland (1999) played excerpts of recordings of speakers from six parts of Wales, as well as standard British speakers, to students and schoolteachers from various parts of Wales. Subjects were asked to identify the dialect of each speaker and to rate the speaker's "Welshness" and likability. The teachers were more accurate at identifying the dialect than the students, but the two groups were fairly consistent in their ratings of speakers' Welshness.

Another experiment involving identification of the dialect of speakers is described in Wolfram, Hazen, and Schilling-Estes (1999, 129–31). They played to listeners four variants of /ɔ/, as in *caught*, uttered by speakers of different dialects. They asked the listeners to rate the variants on scales of most-to-least Northern-sounding and most-to-least Southern-sounding and then to guess where each speaker came from. Their aim was to test reactions to the raised, monophthongal variant of /ɔ/ that occurs around the Pamlico Sound, and the results showed that it ranked high on the Northernness scale and low on the Southernness scale. A rather similar protocol was used by Munro, Derwing, and Flege (1999), who investigated how much the dialect of adults is modified after they move to a new region. Recordings were made of Canadians living in Canada, Canadians living in Alabama, and Alabama natives. Canadian listeners were then asked to rate the voices on a scale from "very Canadian" to "very American," while Alabama subjects were asked to rate the voices on a scale from "definitely from Alabama" to "definitely not from Alabama." Both groups of listeners rated the Canadians living in Alabama as intermediate between the Canadians living in Canada and the Alabama natives, indicating that the Canadi-

ans living in Alabama had undergone some dialectal shifting. The authors determined that the overall rate of speech could not have influenced the ratings, but that diphthong quality may have.

Clopper and Pisoni (2001) investigated how well subjects could distinguish speakers from different parts of the United States and what features they relied on. Sixty-six young, white males from six regions of the United States (11 from each region) read two sentences. These sentences were played to subjects from Indiana, who were asked to match each utterance with a region. Although overall identification accuracy was only 25%, it was greater than chance (17%), and the difference was statistically significant. Comparisons with acoustic measurements of the utterance revealed that the identifications were correlated with certain dialectal variants, including *r*-lessness, /s/ or /z/ in *greasy*, and several aspects of vowel quality.

Reneé van Bezooijen and her colleagues have conducted several dialect identification experiments in the Netherlands and England. Gooskens (1997) and Bezooijen and Gooskens (1999) investigated the relative importance of prosodic and segmental information for listeners reacting to recordings of various dialects. They conducted separate experiments for dialects of Dutch and of British English. Signals to be used for stimuli underwent three treatments. One treatment was low-pass filtering at 400 Hz, which eliminated most of the segmental information in the signals and thus forced listeners to focus on prosody. Another treatment was monotonization, in which the pitch contour of the signals was flattened so that F_0 was always the same, eliminating intonation.[1] The other treatment was simply resynthesis of the original signals to produce control stimuli, which ensured that these stimuli had the same segmental quality as the other two versions. Two different tasks were given to subjects. One was to rate the degree of dialectal divergence of each stimulus from either standard Netherlands Dutch or standard British English. The other was to try to identify what region and locality the speaker of each stimulus was from. The results showed that English listeners relied more on prosody in their judgments of British dialects than Dutch listeners did for Dutch dialects. Bezooijen and Berg (1999) investigated the intelligibility of four regional dialects of Dutch to speakers of near-standard Netherlands Dutch. Speech fragments from the four dialects consisting only of a noun and function words were played to listeners who spoke standard or near-standard Dutch. The listeners were asked to translate the fragments into standard Dutch. Their rates of mistranslations of the nouns were compared with an objective rating of how each noun differed from its equivalent in standard Dutch, that is, the number of different sounds, a different lexical item, or a semantic difference. As expected, the more similar the dialect noun was to

the standard noun, the fewer mistranslations there were. Bezooijen and Ytsma (1999) had listeners identify various Dutch dialects, rate their divergence from each other, and give their subjective personality impressions of them. In general, the more southerly dialects were easiest to identify and rated as most divergent, while standard Dutch was regarded as sounding most arrogant.

IDENTIFICATION OF THE ETHNICITY OF SPEAKERS. Perception experiments testing the ability of listeners to name the ethnicity of speakers have involved playing recordings of African Americans and European Americans, generally either field recordings or tapes of speakers reading a story, to subjects who were asked to name the speakers' ethnicity. Using this method, Dickens and Sawyer (1952); Stroud (1956); Hibler 1960; Larson and Larson (1966); Roberts (1966); Buck (1968); Shuy, Baratz, and Wolfram (1969); Tucker and Lambert (1969); Shuy (1970); Abrams (1973); Irwin (1977); Lass et al. (1979); Bailey and Maynor (1989); Haley (1990); and Trent (1995) found that listeners could identify the ethnicity of speakers much of the time or even nearly all the time. Some three-way ethnic identification studies have been conducted as well. Baugh (1996) showed that listeners could distinguish African Americans, European Americans, and Chicanos; see also Purnell, Idsardi, and Baugh (1999). Wolfram (2000) conducted ethnic identification experiments using speakers from two unusual communities in North Carolina: Hyde County, in which African Americans show considerable assimilation to the local European American dialect, and Robeson County, which has a triethnic assemblage of European Americans, African Americans, and Lumbee Native Americans. Outsiders nearly always misidentified older Hyde County African Americans and usually misidentified Lumbees, but Lumbees themselves were usually able to distinguish other Lumbees. While all of these studies have demonstrated that listeners can often identify the ethnicity of a speaker, their major limitation is that they could not determine what features listeners rely on to make the identifications, though Roberts (1966) found mispronunciations to be associated with identification as African American.

A number of studies have attempted to address this issue. Bryden (1968) conducted two experiments testing whether differences in nasality could serve as a cue. In the first experiment, ten European American and ten African American speakers read a short passage. The recordings were presented either unmodified or with bandwidth compression, which reduced nasality. The ethnicity of speakers was identified accurately more often for the unmodified versions. In the second experiment, 91 speakers of both ethnic groups and a variety of socioeconomic backgrounds read the

same passage and the words *beet* and *boot* in frames. The recordings were labeled for ethnicity and evaluated for speech proficiency. A direct and statistically significant correlation was found between the number of reading errors and perception of a speaker as African American, but correlations with nasality and vowel quality were not significant. Koutstaal and Jackson (1971) determined, by comparing acoustic measurements with identification accuracy, that intonation and timing differences could not be the only cues. Lass, Mertz, and Kimmel (1978) found that playing recordings of African Americans and European Americans backward (which would make the segments difficult to recognize but preserve some aspects of prosody and voice quality) or compressing the recordings temporally reduced the accuracy of ethnic labeling by European American listeners. Lass et al. (1980) played recordings of ten African American and ten European American speakers to listeners under three treatments—unfiltered, low-pass filtered at 225 Hz, and high-pass filtered at 225 Hz—and found that filtering adversely affected the accuracy of ethnic identification by listeners.

Two studies have examined aspects of voice quality more specifically. Hawkins (1993) conducted a two-part study on the role of F_0 in ethnic labeling. In the first phase, equal numbers of African Americans and European Americans, male and female, produced [æ] and [i] in isolation, in words, and in sentences. Listeners of both ethnicities and sexes were able to identify the speaker's ethnicity with better than random accuracy. Measurements of the stimuli implicated F_0 as a cue. In the second phase, isolated [æ] produced by two African American and two European American speakers was synthesized at nine F_0 levels. Listeners equally divided by ethnicity, sex, and residence labeled the ethnicity of the stimuli. Lower F_0 was found to be correlated with labeling as African American, but comparison of the results for different listener groups suggested that the difference was based on stereotype, not physiology. Walton and Orlikoff (1994) presented listeners with 50 paired recordings of sustained /ɑ/ vowels, one spoken by a European American speaker and one by an African American. All 100 of the speakers were adult males. Listeners had to judge which speaker belonged to which ethnicity and were able to do so with 60% accuracy. Then the researchers conducted a careful acoustic analysis of the stimuli, examining jitter (F_0 perturbation), shimmer (amplitude perturbation) and harmonics-to-noise ratio, and found ethnic differences for all three cues. They also found correlations between how much the paired speakers differed in those cues and how accurate the listeners were in their ethnic judgments. In opposition to Hawkins, they speculated that physiology was the basis for the difference.

Other studies have focused on vowels or intonation. Graff, Labov, and Harris (1986) took short excerpts of recordings containing two examples of either /au/ or /o/ and modified F_2 of those vowels in the recordings with a synthesizer. Fronted nuclei of these vowels typify the production of European American Philadelphians and backed nuclei that of African American Philadelphians. African American, European American, and Puerto Rican subjects from Philadelphia were asked to identify the stimuli as African American or European American. The results showed that /au/ or /o/ tokens with higher F_2s were more likely to be identified as European American speech than those with lower F_2s, although the listening task forced subjects to focus on vowel quality, which could easily have biased the results. Purnell, Idsardi, and Baugh (1999) found that European American listeners could distinguish productions of *hello* by African Americans, European Americans, and Chicanos and determined that the quality of /ɛ/ was the most important factor, but that the duration of /ɛ/, the location of the stress on the first or second syllable, and the signal-to-noise ratio also influenced the results. Thomas and Reaser (2002) used five-second excerpts of interviews of Hyde County, North Carolina, natives and control speakers from inland areas. Two excerpts, one featuring diagnostic local vowel variants and the other not doing so, were taken for each Hyde County speaker. The excerpts were given three treatments, much like those in Gooskens (1997): unmodified; monotonized to eliminate intonation; and low-pass filtered at 330 Hz to eliminate most segmental information. The three treatments were played to different groups of listeners. Various prosodic and voice quality factors were measured in the stimuli. For the unmodified and monotonized stimuli, only presence of diagnostic vowels reached or approached statistical significance, and Hyde County African Americans (who exhibit vowel variants more typical of European Americans) were identified less accurately than other groups, indicating that listeners used vowel variation for labeling. For the low-pass filtered stimuli, certain prosodic and voice quality factors reached or approached significance, but only for male voices.

Finally, Foreman (2000) investigated whether intonation could serve as a cue to ethnic identification. African American and European American speakers read scripts that were designed to imitate conversational speech. Recorded sentences were then low-pass filtered at 900 Hz to mask some aspects of segmental and voice quality variation, thereby focusing more attention on intonation. African American and European American listeners were classified according to how much exposure they had to African American and "mainstream" dialects. The results showed that sentences with more diagnostic intonational cues were identified more

accurately than those with fewer intonational cues. In addition, speakers with extensive exposure to both dialects were most accurate in their labeling.

Taken together, these studies give the impression that, under certain circumstances, listeners are capable of accessing a wide variety of cues to determine whether a speaker is African American or European American. Nevertheless, it is not yet clear which ones listeners use in real-life situations or which ones are most important. More studies that compare different cues are needed; single-feature studies have dominated the discussion so far. Additional factors, such as rhythm, should be investigated. In addition, labeling of other ethnic groups, such as Chicanos, and differentiating ethnic groups in other Anglophone countries have received little attention. A study by Heselwood and McChrystal (2000) investigating what features—for example, retroflex stops and clear /l/—distinguish Panjabi English in England represents a start in that direction. More studies of this sort could be conducted on other languages as well. A great deal of work remains to be done on dialect labeling.

IDENTIFICATION OF THE SOCIOECONOMIC LEVEL OF SPEAKERS. Experiments investigating the ability of listeners to name the socioeconomic class of speakers have occasionally been conducted, often in conjunction with ethnic identification. Two early studies by Harms (1961, 1963) demonstrated that listeners of different social strata could often determine the ethnicity of a speaker from a short excerpt. Shuy, Baratz, and Wolfram (1969; also reported in Shuy 1970) had listeners identify the ethnicity and social class of African American and European American speakers from Detroit. They found that lower social classes were identified more accurately than higher ones. Sebastian and Ryan (1985) found that, for both middle-class and lower-class speakers, ratings by listeners for status-related traits were lower for Spanish-accented speakers than for speakers with a "standard" accent. They found in a follow-up experiment that medium-accented speakers were rated higher than low-accented speakers for social class, though high-accented speakers, as expected, were rated lower for social class than other speakers.

INFLUENCE OF STEREOTYPES ON THE PERCEPTION OF SOUNDS. A few studies have examined how stereotypes can alter subjects' perception of speech. Strand (1999) studied how listeners' perception of a speaker's sex affects their identification of sounds. Her study was based in part on the "McGurk effect" (McGurk and MacDonald 1976), in which subjects hearing one sound and watching a video of a speaker uttering a different sound tend to

perceive the sound that they "lip-read" from the video, not the sound that they hear. Strand played to subjects recordings of [sɑ ~ ʃɑ], with the fricative synthesized to match male or female frequencies of [s] or [ʃ]. The vowel had the same formant values for all the stimuli. Subjects simultaneously watched a video of a male or a female uttering such a syllable. The results showed that subjects altered their perception of the fricative depending on the sex of the speaker that they saw, shifting the /s/-/ʃ/ boundary to lower frequencies for male faces and to higher frequencies for female faces. Strand concluded that speech perception is influenced not just by the physical attributes of sound but also by gender stereotypes. The fact that male and female vowel formant values are not scaled uniformly (see, e.g., Fant 1966; Yang 1992) suggests that a similar phenomenon may exist for vowel perception, and, in fact, a follow-up study by Johnson, Strand, and D'Imperio (1999) showed that it does. The latter study examined the /ʊ/-/ʌ/ boundary in *hood* and *hud* and found that it was affected by whether subjects saw, or even imagined, a male or female face saying the words.

Niedzielski (1999) studied a different kind of perceptual shifting caused by stereotyping. It is known that bilinguals shift their perceptual boundaries between sounds depending on which language they believe they are listening to (Elman, Diehl, and Buchwald 1977; Janson and Schulman 1983, 331). Niedzielski found that such shifting may also occur depending on what dialect listeners believe they are hearing or on their expectations of their own speech. She played recordings of /au/, as in *about*, and other vowels spoken by a Detroit native to a group of subjects from the Detroit metropolitan area. Half of the subjects were told that the speaker was from Detroit and the other half that she was from Canada. Subjects were asked to match the vowels that they heard with resynthesized vowels to best approximate the quality of the original vowel. Subjects told that the speaker was from Detroit chose lower /au/ nuclei than those told that the speaker was from Canada, which indicated that their perception was altered by their stereotypes of American and Canadian speech. For the other vowels, they chose qualities that matched widespread American forms more closely than those found in the Detroit dialect. The latter finding apparently resulted from the fact that most Detroit residents do not recognize the distinctiveness of their own speech and hence harbor preconceived notions that their speech is unmarked for dialect features. Both Strand (1999) and Niedzielski (1999) show that speech perception does not depend purely on physical factors, but also on listeners' expectations based on sociological factors. This finding demonstrates that variationists can contribute a great deal to theories of speech perception: it is not just the other way around.

VOWEL MERGERS AND SPLITS IN PERCEPTION. There are a few studies of vowel mergers in perception. Janson and Schulman (1983) used synthetic stimuli to investigate how the perception of residents of Lycksele, Sweden, who distinguish Swedish short /ɛ/ and short /e/ in their production, might differ from that of residents of Stockholm, who merge those vowels in production. Subjects from each community were asked to categorize stimuli that represented a continuum from [a] to [i]. The results showed that most of the Lycksele subjects were unable to distinguish /ɛ/ and /e/ consistently in perception, even though they maintained the distinction in production. The Stockholm subjects, as expected, were unable to distinguish /ɛ/ and /e/. Janson and Schulman suggested that Lycksele speakers had enough exposure to the Stockholm dialect that they had ceased to use the distinction as a means of distinguishing words because it was useless for understanding Stockholm Swedish. Costa and Mattingly (1981) examined /ɑ/ and /ɑr/ in the Boston area, where they are differentiated by length, and found that listeners could not distinguish them. They reached the same conclusion as Janson and Schulman, that distinctions persist in production after they disappear in perception.

Labov, Karen, and Miller (1991) objected to the notion that listeners may cease using a distinction for perception when they retain it in production; they reasoned that the earlier findings were based on experiments involving the labeling of isolated stimuli, an unnatural situation, not on discrimination involving semantic distinctions in running speech. They constructed experiments to test Philadelphians' discrimination abilities with pairs such as *ferry* and *furry*. These pairs are merged by some Philadelphians and in close approximation for many others. The experiments used unsynthesized speech signals. One experiment involved interpretation of an ambiguity in a story, which tested Janson and Schulman's assertion; the other involved identifications of isolated words from commutation tests, in which recorded words uttered by a native speaker of the dialect are played to subjects who are asked to identify them as, for example, *ferry* or *furry*. Their results were not as categorical as those of Janson and Schulman, but they did indicate that many Philadelphians could distinguish nearly merged pairs perceptually, even though they were impaired in their ability to do so. Labov (1994) and Labov and Ash (1997) report results from several other commutation experiments conducted in various locations. Di Paolo and Faber (1990, 166–67) describe a similar test given to subjects in Utah to investigate the near-mergers of /il/ and /ɪl/ (as in *feel* and *fill*, respectively), of /el/ and /ɛl/ (as in *fail* and *fell*), and of /ul/ and /ʊl/ (as in *fool* and *full*).

Perception experiments investigating vowel splits are scarce. Guenter (2000) investigated whether California subjects identified vowels preced-

ing /r/, /l/, and /ŋ/ with vowels in other contexts. He played recordings of a voice saying pairs such as *beer/beet* and *sing/bid*, as well as control pairs such as *grief/beet* and *bit/beet*, and asked the subjects to identify the vowels in the two words as "the same" or "different." He also recorded response times. Subjects showed lower rates of identification as "the same" when one vowel in a word pair preceded /r/, /l/, or /ŋ/ than when identical vowels preceded other consonants. Response times were slower for pairs in which one member preceded /r/, /l/, or /ŋ/ than for pairs in which identical vowels preceded other consonants or pairs in which the vowels differed. The difficulty that subjects exhibited in identifying vowels before /r/, /l/, and /ŋ/ with any phoneme suggested that vowels in these contexts were splitting from their counterparts in other contexts.

DIALECTAL DIFFERENCES IN CATEGORIZATION OF PHONES. Relatively few experimental studies test discrepancies in how speakers of different dialects categorize sounds. Experiments on cross-dialectal differences in the perceptual boundaries between sounds are relatively easy to design and could be applied extensively, though few have been carried out. Willis (1972) asked subjects in Buffalo, New York, and neighboring Fort Erie, Ontario, to categorize synthetic vowels according to phoneme. He found that some of the phoneme boundaries differed for residents of the two communities and that two of the most strongly differing boundaries (/ɛ/-/æ/ and /æ/-/ɑ/) corresponded directly with production differences between the dialects of the communities. Janson (1983, 1986) investigated a difference in generational dialects in Stockholm Swedish, in which /a:/ is undergoing backing. He had subjects identify synthetic stimuli that represented a continuum between the Swedish phrases *ett tag* 'a while' and *ett tåg* 'a train' as one or the other of the two phrases in order to locate the perceptual boundary between /a:/ (as in *tag*) and /o:/ (as in *tåg*). He found that the change was reflected in encroachment of /a:/ on the perceptual space of /o:/ across generations in Stockholm, though the perceptual change was not as fast as the change in production. Furthermore, Janson (1986) found that the shift in the perceptual boundary was not occurring in Helsinki Swedish, which shows no shift in production. Malderez (1995) conducted a similar experiment on the encroachment of /o/ upon /ø/ in French. A potential pitfall of this type of experiment is that, as Niedzielski (1999) found, listeners are not always able to match sounds with their own production norms accurately.

Experiments in cross-dialectal differences in which phonetic cues are used to distinguish sounds are possible as well. In Thomas (2000), I

compared the perception of varying heights of the glide of /ai/, as in *tide* and *tight*, by Anglos from central Ohio and Mexican Americans from southern Texas. In production, both groups produced higher glides before a voiceless consonant, as in *tight*, than before a voiced consonant, as in *tide*, though the Anglos did so to a greater extent than the Mexican Americans. A perception experiment using synthetically modified stimuli representing a continuum from *tide* to *tight* revealed that both groups used the glide difference as a perceptual cue, but in a different way. The Ohio Anglos used the glide difference mainly as a cue to the identity of following [d] versus [t], but the Texas Chicanos used it more as a cue to following [d] versus null (e.g., *tide* vs. *tie*). A possible reason was that the stimuli lacked stop releases, which may have been more important as a cue to the identity of final [t] in Texas Chicano English than in Ohio Anglo English.

A somewhat similar experiment was Peeters's (1991) investigation of cross-linguistic and cross-dialectal differences in steady-state patterns of diphthongs. He generated synthetic stimuli representing [ai], [au], [ei], and [ou] diphthongs but with varying proportions of onset steady state, transition, and offset steady state. Subjects who spoke northern Dutch, British English, dialect-neutral German, and Austrian German were asked to rate the stimuli for goodness—that is, how well they matched their own productions. Each language and dialect investigated differed, indicating that steady-state patterns, like the height of the glide, are a phonetic cue that is ordinarily noncontrastive but can vary dialectally. The method that Peeters employed would be useful for many dialectal comparisons, though the effects of prestige on particular variants would have to be accounted for.

Another type of cross-dialectal perception experiment involves testing how well listeners can identify dialectal pronunciations. William Labov and his team at the University of Pennsylvania have conducted extensive investigations of the perception of geographic dialects using unsynthesized speech samples. Labov, Yaeger, and Steiner (1972, 135–44) describe an experiment involving frontward-gliding forms of /u/ (as in *boot*), /o/ (as in *boat*), and /au/ (as in *out*) uttered by natives of the Pamlico Sound region of North Carolina. Isolated words with these vowels were played to subjects from other regions of the United States, who had almost no success in identifying them correctly. Labov (1994) and Labov and Ash (1997) report results from their Cross-Dialectal Comprehension project. This project used utterances of a variety of vowels by natives of Chicago, Philadelphia, and Birmingham, Alabama (in fact, Labov, Karen, and Miller 1991, discussed above, was a spin-off project). Subjects included listeners from each

of the three cities and from different ethnic groups. Stimuli were gated so that subjects could be asked to identify excised vowels and to identify words excised from their context, with carrier phrases, or in sentences. As expected, subjects had difficulty identifying some of the vowels and words uttered by speakers from cities besides their own, but a rather surprising result was that the listeners sometimes had trouble identifying vowels from their own city.

Similar experiments have been conducted elsewhere. Flanigan and Norris (2000) played words spoken by a southeastern Ohio native that were excised from their context, in phrases, and in sentences to students at different southeastern Ohio campuses, who were asked to identify them. The listeners had a great deal of trouble identifying the words in phrases or in isolation, especially the latter, even though most of them were from southeastern Ohio. Traill, Ball, and Müller (1995) played individual words uttered by speakers of South African English to listeners from the British Isles and asked them to identify each word from three choices. Several of the front vowels, which were undergoing a chain shift in South African English, were prone to misidentifications. A number of experiments have investigated cross-dialectal comprehensibility of running speech, especially comprehensibility of African American speech and comprehension by African Americans of the speech of other groups (e.g., Arahill 1970; Rundell 1973), with a mixture of results.

ASSESSING PERSONAL TRAITS OF SPEAKERS BASED ON VOICES. Experiments in which listeners rate voices for personality traits have been used extensively, largely by social psychologists but also by sociolinguists (e.g., Underwood 1974). Although studies of listeners' assessments of the personality traits of speakers go back as early as Pear (1931), much of the later work has its roots in Lambert et al.'s (1960) matched guise experiment. In this experiment, recordings of bilinguals speaking French or English were played to subjects who rated the speakers on various scales in order to compare their attitudes toward English speakers and French speakers (see also Lambert 1967). The listeners did not recognize that the same individuals were speaking English and French in different stimuli and rated them differently for various traits depending on which language they were speaking. Subjective reaction studies, including matched-guise studies, are too numerous to list here. However, a detailed review of earlier work is found in Giles and Powesland (1975), and reviews of subsequent studies are found in Ryan (1979); Brown and Bradshaw (1985); McMillan and Montgomery (1989, 396–407); Bradac (1990); and Cargile et al. (1994). Ball and Giles (1988) provide guidelines for conducting matched-guise experiments.

Most studies of this sort have used as stimuli voices that were unedited (except for splicing), but a few have employed synthetic modification. One series of these studies began as an attempt to refine the methods of Addington (1968), who had studied the subjective reactions of listeners to utterances by trained speakers varying their voice quality, rate, and pitch. Refinement was desirable because even trained speakers cannot control one vocal factor precisely or without affecting other factors, while a synthesizer can. Brown, Strong, and Rencher (1972, 1973, 1974); Smith et al. (1975); and Apple, Streeter, and Krauss (1979) synthetically modified the intonation and the speaking rate of voices and had listeners rate the resulting stimuli subjectively on scales of benevolence and competence. The relationship between speaking rate and competence was linear, with faster rates being correlated with higher competence. The relationship between speaking rate and benevolence was U-shaped; high and low rates were correlated with low benevolence and moderate rates with high benevolence. There was some controversy about how naturalistic ("ecological") the synthetic stimuli were; see the discussion in Brown, Giles, and Thakerar (1985). Another study that used synthetically modified stimuli is Bezooijen's (1988) investigation of how listeners judge speakers' personalities. She presented listeners in the Netherlands with speech from people of various socioeconomic statuses from the same city (Nijmegen). The speech was presented in four ways: as excerpts of the unaltered recordings; low-pass filtered at 300 Hz, which focused listeners' attention on prosody; with the recordings cut digitally and randomly respliced, which focused listeners' attention on voice quality; and as written text. The listeners rated the speech on various personality scales—"much education," "strong-willed," and "fair." Then the original recordings were rated on voice quality scales by a different set of listeners, and statistical comparisons between the personality ratings and the voice quality ratings were run. The correlations suggested that listeners inferred a strong personality from prosody, but that they inferred a speaker's intellectual qualities and socioeconomic status from segmental aspects of speech.

Fewer experiments have involved evaluations of voices for traits other than personality features. Examples of such studies are Labov (1966), Frazer (1987), and Plichta (2001). Labov assembled stimuli consisting of 22 sentences uttered by five women from the Lower East Side of Manhattan and played them to other residents of the Lower East Side. The stimuli highlighted several phonological variables that were important in New York City. Subjects were asked to rate the speakers' suitability for various jobs. In general, speakers who themselves showed a high degree of stylistic conditioning of a local variant tended to be most aware of it, as indicated by

negative responses upon hearing it. Frazer (1987) played excerpts of the reading of a story by rural Illinois natives, highlighting several variables. Listeners, who were college students from Illinois, were asked to rate the excerpts according to the question, "Would you be proud if a friend or member of your family spoke this way?" The results showed that stimuli exhibiting variants associated with Southernness were evaluated negatively. Plichta (2001) showed European American and African American subjects videotapes of two African Americans and two European Americans reading a passage; the audio components of the videos were interchanged. Subjects were asked to rate the speakers they saw according to standardness, similarity to their own speech, education, and region. The most important result was that, although European American subjects rated speakers of each ethnicity equally highly for standardness, African American subjects tended to rate African American speakers lower for standardness than they rated European American speakers.

GUIDELINES FOR PERCEPTION EXPERIMENTS

As the studies discussed above indicate, sociolinguists have a wide variety of experimental issues and designs from which to choose for perception experiments. Preparation is the most difficult part of a perception study. The following considerations offer a cookbook approach to conducting perception experiments.

CHOICE OF SPEAKERS. For the vast majority of socioperceptual experiments, real voices, rather than artificially generated ones (text-to-speech), are desirable as the basis of the stimuli, even if the voice will be synthetically modified. As a result, it is important to consider carefully what speakers are chosen and what sort of speech they are asked to produce. Many experiments require read speech, but reading fluency varies among speakers and may become a source of spurious variation. Excluding speakers who are not fluent readers may produce a bias for other variables, however. Conversational interviews may provide the most naturalistic speech, but the content cannot be controlled and sound quality may be defective, especially for field interviews. Another issue is how strongly marked a speaker's dialect is. Although it may be advantageous to use speakers who show traits of the dialect under study especially strongly, such speakers may not typify their speech communities, which could add a bias to the study. Conversely, speakers who are most easily accessible to researchers, such as university students, may be more exposed to outside influences and may not be

especially vernacular because of their socioeconomic background. In, for example, a dialect identification experiment, these sorts of biases could cause listeners to rely on different variants than they ordinarily would or to emphasize certain cues more than usual. If a stereotyped variant occurred in a stimulus, it would quite likely divert listeners' attention away from other cues. While these issues may not affect every experiment, they should be considered by experimenters.

RECORDING EQUIPMENT AND CONDITIONS FOR SIGNALS TO BE MODIFIED. The sound quality of the recordings to be used as stimuli is crucial, especially if the recordings are to be modified with a synthesizer, because a low signal-to-noise ratio in the recordings—caused by background noise, machine noise, or poor reception of the speaker's voice—can make synthesis programs inoperable. For this reason, field recordings are often unusable or difficult to use with synthesizers, though it may be worthwhile to try anyway. Even if the stimuli are not synthetically altered, recording problems and variations in recording quality can affect the responses of subjects. When possible, recordings to be synthetically modified should be made in a soundproof or sound-treated room with a directional microphone. If such a room is unavailable, however, recordings made in a quiet room may be the best option. Recording conditions should be noted.

Researchers usually report the make and model of recorders, microphones, sound rooms, and such as a matter of course. The make of equipment can be a factor, but, obviously, factors such as age, wear, and cleanliness can also affect equipment performance. For that reason some editors ask that researchers perform a frequency response test on equipment, especially earphones. Frequency response tests determine how well the equipment handles various frequencies. Unfortunately, the equipment needed for frequency response tests is often unavailable to sociolinguists. Nevertheless, equipment deterioration primarily affects high-frequency sounds such as sibilants, so it is normally a minimal factor for most of the variables that sociolinguists study. For that reason, frequency response tests may not be crucial for the majority of sociophonetic experiments. A comparison of some currently available models of microphones and tape recorders is given in Plichta and Mendoza-Denton (2001).

TYPES OF SYNTHESIS AND SYNTHESIZERS. Synthetic stimuli are not necessary for all perception experiments, and for some they are not even desirable (e.g., Labov, Karen, and Miller 1991). However, the speech synthesizer is a valuable tool that sociolinguists have underutilized, and it is indispensable

for some types of experiments. A brief description of synthesis issues relevant to sociolinguists is given here.

Most synthesis packages available today can modify signals that are fed into them. Many synthesizers also have text-to-speech, that is, a dictionary of individual words that they can generate and the ability to string words together into phrases and sentences. For sociolinguists, feeding signals into the synthesizer is normally more useful than text-to-speech because the input will be naturalistic. Words that are generated by the synthesizer depend on the comprehensiveness of the synthesis program, and since there are aspects of speech that are not completely understood, no synthesis program is perfect. The problems are especially acute for strings of several words. For any sociolinguistic study involving text-to-speech, it may be necessary for listeners to rate the tokens for naturalness.

The two types of synthesizers seen most often today are LPC synthesizers and Klatt synthesizers. LPC synthesizers, which are analysis-synthesis systems, are simpler in their design. Their main limitation is that LPC (linear predictive coding) is useful only for periodic sounds, so LPC synthesizers are poorly suited for modifying frication noise or aspiration. Klatt synthesizers, which fall within a group called terminal analog formant synthesizers, can produce aperiodic as well as periodic sounds. Production of vowels and approximants can be set in either cascade mode, in which various filters are employed sequentially, or in parallel mode, in which the filters are employed simultaneously. Consonantal sounds are produced in parallel mode. The cascade mode is used more often for vowels, because it sets formant amplitudes automatically. Klatt synthesizers were designed for text-to-speech, but some of the newer packages that contain a Klatt synthesis system also include capabilities for modifying signals. For more detailed discussions and descriptions of other kinds of synthesizers, see Klatt (1987) and Carlson and Granström (1997).

One problem that can be encountered when only part of a running signal is synthetically modified or when different signals are spliced together is that bursts of noise, or transience, can result at boundaries between unmodified and modified sections or at boundaries between spliced sections. Transience may be quite noticeable and distracting to listeners. It can be avoided by starting and ending modified sections at zero-crossing points, that is, points where the waveform crosses the zero amplitude mark. Another remedy for this problem is to attenuate the signal gradually toward boundaries.

Various synthetic treatments can be performed on stimuli. Low-pass filtering is often used to focus subjects' attention on prosody. Monotonization

is used to eliminate intonation. The fundamental frequency can be modified in less drastic ways as well, such as for modifying intonational contours or for changing voice quality. Various mutations of temporal factors can be used to change the rhythm or to affect the identity of segments, particularly for voiced versus voiceless consonants. Formant values can be changed to constant values, such as those for schwa, to eliminate vowel quality information. Formant values can also be changed to modify particular segments—for example, to make them sound like variants found in a different dialect or like a different phoneme. Splicing and random reordering of portions of the signal are occasionally used to focus subjects' attention on voice quality. The stimuli resulting from these operations should be checked auditorily, because one of the hazards of such modifications is that the synthesized tokens may not always represent accurately what they are supposed to represent. In addition, as Bezooijen and Boves (1986) note, some standard synthetic operations do not accomplish what they are commonly assumed to do. They found that low-pass filtering preserves voice quality factors as well as prosody and that splicing and random reordering do not eliminate all aspects of segmental quality or prosody.

There are also some other hazards with the use of synthesized speech that have to do with experimental design, since synthetic stimuli are usually played to subjects under controlled conditions. The first is that the content of the stimuli can sometimes force listeners to pay more attention to a particular perceptual cue than they ordinarily would, especially if other cues that they usually rely on are absent. The second hazard, described by Labov, Karen, and Miller (1991, 52–53), is that social norms biased against variants represented in the stimuli can skew reactions from subjects. The third hazard, mentioned in Labov (1994, 402), has to do with the fact that subjects in experiments involving synthetic stimuli are usually asked to label the stimuli as, for example, one word or another. At issue is the fact that labeling, a mostly conscious act, is not the same as perception in conversations, which is largely a subconscious act.

In spite of these pitfalls, perception experiments involving synthetic stimuli—when used judiciously—allow researchers to impose tighter experimental controls than experiments using recordings of real speech. For example, with a synthesizer, an experimenter can be certain that a phonetic factor is being modified at steady increments and independently of other factors, something that is impossible with even an adept impersonator. Synthesizers can also provide data on some issues that no other technique can. They are most useful for modifying vowel formants, the fundamental frequency, and the duration of parts of the signal, and can be used for some

consonantal attributes, such as voice onset time. Though sociolinguists have used synthesizers sparingly, they represent perhaps the biggest single aid in perception research.

LISTENING EQUIPMENT AND ENVIRONMENT. The factors noted above for recording equipment also apply to listening equipment; the main difference is that the use of earphones is an issue. Phoneticians usually prefer to play the stimuli to subjects through earphones. Earphones are useful for some types of experiments, especially when the experiment involves a small number of listeners. The use of earphones, of course, creates a sociolinguistically unnatural environment, so many sociolinguists may prefer not to employ them. In addition, earphones may be impractical for large numbers of subjects or when subjects are recruited in the field. It may instead be more expedient to play stimuli to subjects in a quiet location, such as a classroom. This approach is useful when the experiment can be linked to class discussions of phonetics or sociolinguistics.

For tests conducted in classrooms, any loudspeaker system in good working order is acceptable, but researchers should remember that high-frequency sounds travel better forward than sideways, while low-frequency sounds travel about equally well in all directions. This problem mainly affects sibilants, but it is still important to note the placement of the loudspeakers relative to subjects and perhaps how the subjects are seated in the room. For any perception test conducted without earphones, whether in a classroom or in the field, it is crucial to eliminate as much background noise as possible. When earphones are not used, it may be important to take note of the acoustic reflectivity, that is, the amount of echo, in the listening environment. The volume should be set to a level comfortable for subjects, and either the test administrator or the subjects may adjust the volume.

SCREENING FOR HEARING IMPAIRMENT. Subjects in perception tests should be asked whether they have any hearing impairment. Those who report hearing impairment should ordinarily be excluded from the analysis, though for special groups of listeners, such as elderly subjects, it may be necessary to include subjects with some hearing loss. Speech pathologists and phoneticians often conduct their own hearing screenings before conducting a perception test, but doing so is likely to be impractical for most sociolinguists. Inquiring whether any subjects have sinus or ear infections, which could affect their hearing, may be advisable.

EXPERIMENTAL TASK. It almost goes without saying that the task that subjects are asked to perform should be tailored to the issue being investigated. Common tasks for subjects in perception experiments include (1) judging whether two stimuli sound the same or different; (2) judging which two stimuli out of three or more are alike; (3) selecting which phoneme, word, ethnic group, or such out of two or more choices a stimulus matches most closely; (4) identifying a stimulus open-endedly as a particular word, or such, and writing it down; and (5) gauging how realistic, natural, or typical of a particular group a stimulus sounds. Of course, researchers can dream up any number of other possible tasks.

For some types of experiments, it may be important to take the sociolinguistic environment into consideration. Lee (1971) discusses this problem in detail, noting that experiments with carefully controlled stimuli may not resemble natural listening conditions, especially when the discourse context is an important factor. He also notes various other problems, such as instances in which the variables being tested are transparent to subjects or other uncontrolled variables influence responses. The last problem may be remedied by synthetic manipulation. Nevertheless, Lee's criticisms point out the importance of careful experimental design and of recognizing the limitations of any experiment.

PRESENTATION OF STIMULI. The way that stimuli are presented to subjects should be noted. For the most part, this means describing the length of pauses between stimuli, the number of stimuli per set, how subjects were cued that a set of stimuli had ended, whether the subjects simply listened to a recording or interacted with the person administering the test, and the total length of each trial. For most tasks, the order of stimuli should be randomized. Usually, trials should last no longer than 10–15 minutes because fatigue by subjects sets in at that point. In addition, it often takes subjects a little while to become adjusted to a task, so—if a long series of stimuli is presented—results from the first few stimuli that subjects hear should be discarded. These stimuli can be repeated later during the trial. The number of stimuli from the beginning that are discarded should be reported.

RESPONSE FORM. The general manner in which subjects give responses should be described. The traditional method is to use an answer sheet, but computers may be more convenient if subjects are being tested one at a time, and laptop computers may be particularly convenient in the field. For certain experiments, subjects may even say their responses aloud while the

test administrator writes them down or tape-records them. Some experiments record not only the responses of subjects but their response time as well. When response time is recorded, it is important to note whether the time includes the time it took to play the stimulus.

FURTHER DIRECTIONS

The studies reviewed earlier illustrate the variety of research questions in language variation to which perception experiments can be applied. Nevertheless, there is considerable room for expansion. Several of the topics discussed above, such as the effects of gender stereotypes on perception, have scarcely been touched. Furthermore, numerous other issues remain for which perception experiments could be useful.

One such issue is the accuracy of impressionistic phonetic transcription, of which some aspects have been examined (e.g., Ladefoged 1960; Kerswill and Wright 1990; Nairn and Hurford 1995; Heuvel and Cucchiarini 2001). Both dialectologists and (as noted earlier) sociolinguists have traditionally relied heavily on impressionistic transcription in spite of its inherent subjectivity. Other problems are associated with it as well. The transcriptions may be difficult to interpret (see, e.g., McDavid 1981), especially when different transcribers are involved (see the discussion of "field-worker isoglosses" in Trudgill 1983, 38–41; see also Allen 1976, 23). A related problem is that, as Trudgill (1983, 35–38) and Labov (1994, 74) note, transcribers' preconceptions affect their transcriptions: old, well-known variants are differentiated extensively in the transcriptions, while newer, lesser-known variants tend to be undertranscribed or ignored altogether. Perception experiments could examine the extent to which different transcribers are affected by preconceptions and by the influence of their own native dialects. A different sort of problem has to do with the fact that listeners compensate perceptually for processes that affect sounds in running speech, including the coarticulatory effects of neighboring segments (Lindblom and Studdert-Kennedy 1967; Ohala 1981, 179–87; Ohala and Feder 1994; Nábělek and Ovchinnikov 1997; Holt, Lotto, and Kluender 2000), the effects of durational variation (Lindblom and Studdert-Kennedy 1967; Janson 1979), and, probably, the effects of variation in stress. In all likelihood, transcribers are affected by this perceptual compensation and, as a result, may be unable to detect many effects of phonetic context. Of course, transcribers can hear some contextual differences. Perception experiments could shed more light on the degree to which transcribers are able to escape their own perceptual compensation.

This perceptual compensation, which Ohala (1981, 183) calls "reconstructive rules," could potentially open up another new line of perception experiments for variationists. The studies noted above that have demonstrated that compensation occurs avoided the issue of dialectal variation (except insofar as Ohala 1981 relates it to diachronic variation). One may assume that listeners' reconstructive rules are oriented toward varieties of language most similar to the listeners' own speech. What happens when listeners are subjected to unfamiliar dialects, however? Do the listeners' reconstructive rules lead to errors in perception? Moreover, are listeners who have extensive exposure to a different dialect able to shift to a perceptual mode appropriate to that dialect? Another question is whether listeners exhibit different compensatory patterns for different styles of speech, given that speakers may adjust how carefully they articulate depending on speech style (Lindblom 1990). Variationists could address these questions experimentally.

Another perceptual issue to which variationists could contribute is the controversy between what Strange (1989) calls the "elaborated target model" and "dynamic specification" theories of vowel recognition; this issue has attracted a great deal of attention from experimental phoneticians recently. The elaborated target model is based on the notion that listeners recognize vowels from steady states, while the dynamic specification model states that listeners recognize vowels from the formant movement found in consonant transitions. Yet another model states that diphthongal formant movement is an important cue; see the discussion of all three models in Pitermann (2000). Target models in various forms are the oldest theory, and there is evidence that at least some listeners rely on steady states (Gottfried 1984; Pitermann 2000; Sussman 2001). However, there is also considerable evidence, in large part from experiments in which the center of vowels is gated out of signals, that listeners can recognize vowels from consonantal transitions (Verbrugge and Rakerd 1986; Strange 1989; Strange and Bohn 1998) or diphthongal formant movement (Assman, Nearey, and Hogan 1982; Andruski and Nearey 1992; Hillenbrand and Nearey 1999). The major questions are which type of recognition is primary and which one listeners generally rely upon in natural dialogue. Nearey (1989) suggested that listeners may use all of these strategies, depending on the situation. Gottfried (1984) reported that French speakers do not rely on dynamic properties of a vowel when listening to French, unlike English speakers listening to English, suggesting that a target model strategy is more usual for French. He speculated that the reason was that French vowels are monophthongal, whereas the gliding that typifies most varieties of English promotes strategies that focus on dynamic properties of vowels.

His finding suggests that variationists could test the models of vowel recognition by comparing the behavior of speakers of mainstream dialects of English with those showing little or no gliding, for example, varieties with considerable substratum effects from other languages. Other experiments could test how the different strategies operate in running speech, perhaps by gating out parts of a vowel spoken in a context in order to create the potential for ambiguity.

Even though some of the topics mentioned in this section and earlier are usually studied by nonsociolinguists, especially phoneticians and psychologists, variationists should not shy away from them. Cognizance of intraspeaker linguistic differences (i.e., stylistic and register variation) and of various interpersonal differences (e.g., those related to geography, ethnicity, gender, and socioeconomic status) is fundamental to sociolinguistics, but these differences are often regarded lightly by the more experimental branches of linguistics. Lee (1971) notes that one of the chief shortcomings of many perceptual studies of dialectal variation is the failure to take the sociolinguistic setting into account, yet it is this shortcoming that sociolinguists are best equipped to address. An approach to speech perception that not only takes variation into account, but also recognizes its usefulness for experimental design, can advance the understanding of numerous aspects of speech perception greatly. Furthermore, if sociolinguists devote more attention to perception, their understanding of language variation will deepen significantly.

NOTES

I wish to thank Paul Foulkes, John J. Ohala, and two anonymous reviewers for their comments and suggestions regarding this project and Walt Wolfram for his support. Research for this paper was supported in part by the National Science Foundation grant BCS 99-10224.

1. The fundamental frequency (F_0) represents the rate at which the vocal folds vibrate. Formants (F_1, F_2, etc.) are resonances produced by the vocal tract. Both are normally measured in cycles per second, or Hertz (Hz).

REFERENCES

Abrams, Albert S. 1973. "Minimal Auditory Cues for Distinguishing Black from White Talkers." Ph.D. diss., City Univ. of New York.

Addington, David W. 1968. "The Relationship of Selected Vocal Characteristics to Personality Perception." *Speech Monographs* 35: 492–503.

Allen, Harold B. 1976. *The Linguistic Atlas of the Upper Midwest.* Vol. 3, *The Pronunciation.* Minneapolis: Univ. of Minnesota Press.

Andruski, Jean E., and Terrance M. Nearey. 1992. "On the Sufficiency of Compound Target Specification of Isolated Vowels and Vowels in /bVb/ Syllables." *Journal of the Acoustical Society of America* 91: 390–410.

Apple, William, Lynn A. Streeter, and Robert M. Krauss. 1979. "Effects of Pitch and Speech Rate on Personal Attributions." *Journal of Personality and Social Psychology* 37: 715–27.

Arahill, Edward Joseph. 1970. "The Effect of Differing Dialects upon the Comprehension and Attitude of Eighth Grade Children." Ph.D. diss., Univ. of Florida.

Assman, Peter F., Terrance M. Nearey, and John T. Hogan. 1982. "Vowel Identification: Orthographic, Perceptual, and Acoustic Aspects." *Journal of the Acoustical Society of America* 71: 975–89.

Bailey, Guy, and Natalie Maynor. 1989. "The Divergence Controversy." *American Speech* 64: 12–39.

Ball, Peter, and Howard Giles. 1988. "Speech Style and Employment Selection: The Matched Guise Technique." In *Doing Social Psychology: Laboratory and Field Exercises*, ed. Glynis M. Breakwell, Hugh Foot, and Robin Gilmour, 121–49. Cambridge: Cambridge Univ. Press.

Baugh, John. 1996. "Perceptions within a Variable Paradigm: Black and White Racial Detection and Identification Based on Speech." In *Focus on the USA*, ed. Edgar W. Schneider, 169–82. Amsterdam: Benjamins.

Bezooijen, Renée van. 1988. "The Relative Importance of Pronunciation, Prosody and Voice Quality for the Attribution of Social Status and Personality Characteristics." In *Language Attitudes in the Dutch Language Area*, ed. Roeland van Hout and Uus Knops, 85–103. Dordrecht: Foris.

Bezooijen, Renée van, and Rob van den Berg. 1999. "Word Intelligibility of Language Varieties in the Netherlands and Flanders under Minimal Conditions." In *Linguistics in the Netherlands 1999*, ed. René Kager and Renée van Bezooijen, 1–12. Amsterdam: Benjamins.

Bezooijen, Renée van, and Louis Boves. 1986. "The Effects of Low-Pass Filtering and Random Splicing on the Perception of Speech." *Journal of Psycholinguistic Research* 15: 403–17.

Bezooijen, Renée van, and Charlotte Gooskens. 1999. "Identification of Language Varieties: The Contribution of Different Linguistic Levels." *Journal of Language and Social Psychology* 18: 31–48.

Bezooijen, Renée van, and Jehannes Ytsma. 1999. "Accents of Dutch: Personality Impression, Divergence, and Identifiability." *Belgian Journal of Linguistics* 13: 105–29.

Bradac, James J. 1990. "Language Attitudes and Impression Formation." In *Handbook of Language and Social Psychology*, ed. Howard Giles and W. Peter Robinson, 387–412. New York: Wiley.

Brennan, Eileen M., and John S. Brennan. 1981. "Measurements of Accent and Attitude toward Mexican-American Speech." *Journal of Psycholinguistic Research* 10: 487–501.

Brennan, Eileen M., Ellen Bouchard Ryan, and William E. Dawson. 1975. "Scaling of Apparent Accentedness by Magnitude Estimation and Sensory Modality Matching." *Journal of Psycholinguistic Research* 4: 27–36.

Browman, Catherine P., and Louis Goldstein. 1990. "Tiers in Articulatory Phonology, with Some Implications for Casual Speech." In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, ed. John Kingston and Mary Beckman, 341–76. Cambridge: Cambridge Univ. Press.

———. 1991. "Gestural Structures: Distinctiveness, Phonological Processes, and Historical Change." In *Modularity and the Motor Theory of Speech Perception: Proceedings of a Conference to Honor Alvin M. Liberman*, ed. Ignatius G. Mattingly and Michael Studdert-Kennedy, 313–38. Hillsdale, N.J.: Erlbaum.

Brown, Bruce L., and Jeffrey M. Bradshaw. 1985. "Towards a Social Psychology of Voice Variations." In *Recent Advances in Language, Communication, and Social Psychology*, ed. Howard Giles and Robert N. St. Clair, 144–81. London: Erlbaum.

Brown, Bruce L., Howard Giles, and Jitendra N. Thakerar. 1985. "Speaker Evaluations as a Function of Speech Rate, Accent, and Context." *Language and Communication* 5: 207–20.

Brown, Bruce L., William J. Strong, and Alvin C. Rencher. 1972. "Acoustic Determinants of Perceptions of Personality from Speech." *International Journal of the Sociology of Language* 6: 11–32.

———. 1973. "Perceptions of Personality from Speech: Effects of Manipulations of Acoustical Parameters." *Journal of the Acoustical Society of America* 54: 29–35.

———. 1974. "Fifty-four Voices from Two: The Effects of Simultaneous Manipulations of Rate, Mean Fundamental Frequency and Variance of Fundamental Frequency on Ratings of Personality from Speech." *Journal of the Acoustical Society of America* 55: 313–18.

Bryden, James D. 1968. *An Acoustic and Social Dialect Analysis of Perceptual Variables in Listener Identification and Rating of Negro Speakers*. ED 022 186. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare, Office of Education, Bureau of Research.

Buck, Joyce F. 1968. "The Effects of Negro and White Dialectal Variations upon Attitudes of College Students." *Speech Monographs* 35: 181–86.

Bush, Clara N. 1967. "Some Acoustic Parameters of Speech and Their Relationships to the Perception of Dialect Differences." *TESOL Quarterly* 1.3: 20–30.

Cargile, Aaron C., Howard Giles, Ellen B. Ryan, and James J. Bradac. 1994. "Language Attitudes as a Social Process: A Conceptual Model and New Directions." *Language and Communication* 14: 211–36.

Carlson, Rolf, and Björn Granström. 1997. "Speech Synthesis." In *The Handbook of Phonetic Sciences*, ed. William J. Hardcastle and John Laver, 768–88. Oxford: Blackwell.

Clopper, Cynthia G., and David B. Pisoni. 2001. "Some Acoustic Cues for Categorizing American English Regional Dialects." Paper presented at the 30th conference on New Ways of Analyzing Variation (NWAV 30), Raleigh, N.C., 11–14 Oct.

Costa, Paul, and Ignatius G. Mattingly. 1981. "Production and Perception of Phonetic Contrast during Phonetic Change." *Journal of the Acoustical Society of America* 69: S67.

Di Paolo, Marianna, and Alice Faber. 1990. "Phonation Differences and the Phonetic Content of the Tense-Lax Contrast in Utah English." *Language Variation and Change* 2: 155–204.

Dickens, Milton, and Granville M. Sawyer. 1952. "An Experimental Comparison of Vocal Quality among Mixed Groups of Whites and Negroes." *Southern Speech Journal* 17: 178–85.

Elman, Jeffrey L., Randy L. Diehl, and Susan E. Buchwald. 1977. "Perceptual Switching in Bilinguals." *Journal of the Acoustical Society of America* 62: 971–74.

Fant, Gunnar. 1966. "A Note on Vocal Tract Size Factors and Non-uniform F-Pattern Scalings." *Speech Transmission Laboratory—Quarterly Progress and Status Report (STL-QPSR)* 4: 22–30.

Flanigan, Beverly Olson, and Franklin Paul Norris. 2000. "Cross-Dialectal Comprehension as Evidence for Boundary Mapping: Perceptions of the Speech of Southeastern Ohio." *Language Variation and Change* 12: 175–201.

Foreman, Christina. 2000. "Identification of African-American English from Prosodic Cues." *Texas Linguistic Forum* 43: 57–66.

Foulkes, Paul. 1997. "Historical Laboratory Phonology—Investigating /p/ > /f/ > /h/ Changes." *Language and Speech* 40: 249–76.

Frazer, Timothy C. 1987. "Attitudes toward Regional Pronunciation." *Journal of English Linguistics* 20: 89–100.

Giles, Howard, and Peter F. Powesland. 1975. *Speech Style and Social Evaluation*. New York: Academic.

Goldbeck, T. P., R. Standke, and K. R. Scherer. 1988. "Digital Techniques for Signal Processing in Vocal Communication Research." *Psychologische Rundschau* 39: 191–200.

Gooskens, Charlotte. 1997. "On the Role of Prosodic and Verbal Information in the Perception of Dutch and English Language Varieties." Ph.D. diss., Katholieke Universiteit Nijmegen.

Gottfried, Terry L. 1984. "Effect of Consonant Context on the Perception of French Vowels." *Journal of Phonetics* 12: 91–114.

Graff, David, William Labov, and Wendell A. Harris. 1986. "Testing Listeners' Reactions to Phonological Markers of Ethnic Identity: A New Method for Sociolinguistic Research." In *Diversity and Diachrony*, ed. David Sankoff, 45–58. Amsterdam Studies in the Theory and History of Linguistic Science, ser. 4: Current Issues in Linguistic Theory 53. Amsterdam: Benjamins.

Guenter, Joshua. 2000. "Vowels of California English before /r/, /l/, and /ŋ/." Ph.D. diss., Univ. of California at Berkeley.

Guy, Gregory R. 1980. "Variation in the Group and the Individual: The Case of Final Stop Deletion." In *Locating Language in Time and Space*, ed. William Labov, 1–36. Orlando: Academic.

Haley, Kenneth. 1990. "Some Complexities of Speech Identification." *SECOL Review* 14: 101–13.

Harms, L. S. 1961. "Listener Judgments of Status Cues in Speech." *Quarterly Journal of Speech* 47: 164–68.

———. 1963. "Status Cues in Speech." *Lingua* 12: 300–306.

Hawkins, Francine Dove. 1993. "Speaker Ethnic Identification: The Roles of Speech Sample, Fundamental Frequency, Speaker and Listener Variations." Ph.D. diss., Univ. of Maryland at College Park.

Heselwood, Barry, and Louise McChrystal. 2000. "Gender, Accent Features and Voicing in Panjabi-English Bilingual Children." *Leeds Working Papers in Linguistics* 8: 45–70.

Heuvel, Henk van den, and Catia Cucchiarini. 2001. "/r/-Deletion in Dutch: Rumours or Reality?" In *'r-atics: Sociolinguistic, Phonetic and Phonological Characteristics of* /r/, ed. Hans Van de Velde and Roeland van Hout, 185–98. Etudes et Travaux. Brussels: Institut des Langues Vivantes et de Phonétique, Université Libre de Bruxelles.

Hibler, Madge Beatrice. 1960. "A Comparative Study of Speech Patterns of Selected Negro and White Kindergarten Children." Ph.D. diss., Univ. of Maryland at College Park.

Hillenbrand, James M., and Terrance M. Nearey. 1999. "Identification of Resynthesized /hVd/ Utterances: Effects of Formant Contour." *Journal of the Acoustical Society of America* 105: 3509–23.

Holt, Lori L., Andrew J. Lotto, and Keith R. Kluender. 2000. "Neighboring Spectral Content Influences Vowel Identification." *Journal of the Acoustical Society of America* 108: 710–22.

Hombert, Jean-Marie, John J. Ohala, and William G. Ewan. 1979. "Phonetic Explanations for the Development of Tones." *Language* 55: 37–58.

Irwin, Ruth Beckey. 1977. "Judgments of Vocal Quality, Speech Fluency, and Confidence of Southern Black and White Speakers." *Language and Speech* 20: 261–66.

Janson, Tore. 1979. "Vowel Duration, Vowel Quality, and Perceptual Compensation." *Journal of Phonetics* 7: 93–103.

———. 1983. "Sound Change in Perception and Production." *Language* 59: 18–34.

———. 1986. "Sound Change in Perception: An Experiment." In *Experimental Phonology*, ed. John J. Ohala and Jeri J. Jaeger, 253–60. Orlando: Academic.

Janson, Tore, and Richard Schulman. 1983. "Non-distinctive Features and Their Use." *Journal of Linguistics* 19: 321–36.

Johnson, Keith, Elizabeth A. Strand, and Mariapaola D'Imperio. 1999. "Auditory-Visual Integration of Talker Gender in Vowel Perception." *Journal of Phonetics* 27: 359–84.

Jonasson, Jan. 1971. "Perceptual Similarity and Articulatory Reinterpretation as a Source of Phonological Innovation." *Papers from the Institute of Linguistics, University of Stockholm* 8: 30–42.

Kerswill, Paul, and Susan Wright. 1990. "The Validity of Phonetic Transcription: Limitations of a Sociolinguistic Research Tool." *Language Variation and Change* 2: 255–75.

Klatt, Dennis H. 1987. "Review of Text-to-Speech Conversion for English." *Journal of the Acoustical Society of America* 82: 737–93.

Koutstaal, Cornelis W., and Faith L. Jackson. 1971. "Race Identification on the Basis of Biased Speech Samples." *Ohio Journal of Speech and Hearing* 6: 48–51.

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

———. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: Univ. of Pennsylvania Press.

———. 1994. *Principles of Linguistic Change*. Vol. 1, *Internal Factors*. Language in Society 20. Oxford: Blackwell.

Labov, William, and Sharon Ash. 1997. "Understanding Birmingham." In *Language Variety in the South Revisited*, ed. Cynthia Bernstein, Thomas Nunnally, and Robin Sabino, 508–73. Tuscaloosa: Univ. of Alabama Press.

Labov, William, Mark Karen, and Corey Miller. 1991. "Near-Mergers and the Suspension of Phonemic Contrast." *Language Variation and Change* 3: 33–74.

Labov, William, Malcah Yaeger, and Richard Steiner. 1972. *A Quantitative Study of Sound Change in Progress*. Philadelphia: U.S. Regional Survey.

Ladefoged, Peter. 1960. "The Value of Phonetic Statements." *Language* 36: 387–96.

Lambert, Wallace E. 1967. "A Social Psychology of Bilingualism." *Journal of Social Issues* 23: 91–108.

Lambert, W[allace] E., R. C. Hodgsen, R. D. Gardner, and S. Fillenbaum. 1960. "Evaluational Reaction to Spoken Language." *Journal of Abnormal and Social Psychology* 60: 44–51.

Larson, Vernon S., and Carolyn H. Larson. 1966. "Reactions to Pronunciation." In *Communication Barriers for the Culturally Deprived*, ed. Raven I. McDavid, Jr., and William M. Austin. Cooperative Research Project No. 2107. Washington, D.C.: U.S. Office of Education.

Lass, Norman J., Celest A. Almerino, Laurie F. Jordan, and Jayne M. Walsh. 1980. "The Effect of Filtered Speech on Speaker Race and Sex Identifications." *Journal of Phonetics* 8: 101–12.

Lass, Norman J., Pamela J. Mertz, and Karen L. Kimmel. 1978. "The Effect of Temporal Speech Alterations on Speaker Race and Sex Identifications." *Language and Speech* 21: 279–90.

Lass, Norman J., John E. Tecca, Robert A. Mancuso, and Wanda I. Black. 1979. "The Effect of Phonetic Complexity on Speaker Race and Sex Identifications." *Journal of Phonetics* 7: 105–18.

Lee, Richard R. 1971. "Dialect Perception: A Critical Review and Re-evaluation." *Quarterly Journal of Speech* 57: 410–17.

Lindblom, Björn. 1980. "The Goal of Phonetics, Its Unification and Application." *Phonetica* 37: 7–26.

———. 1990. "Explaining Phonetic Variation: A Sketch of the H&H Theory." In *Speech Production and Speech Modelling*, ed. W. J. Hardcastle and A. Marchal, 403–39. Dordrecht: Kluwer.

Lindblom, Björn, and Michael Studdert-Kennedy. 1967. "On the Role of Formant Transitions in Vowel Recognition." *Journal of the Acoustical Society of America* 42: 830–43.

Malderez, Isabelle. 1995. "The Use of a Category-Perception Test in the Study of Ongoing Sound Change." In *Proceedings of the 13th International Congress of Phonetic Sciences, ICPhS 95: Stockholm, Sweden, 13–19 August 1995*, ed. Kjell Elenius and Peter Branderud, 684–87. Stockholm: Royal Institute of Technology and Stockholm Univ.

McDavid, Raven I., Jr. 1981. "Low-Back Vowels in Providence: A Note in Structural Dialectology." *Journal of English Linguistics* 15: 21–29.

McGurk, Harry, and John MacDonald. 1976. "Hearing Lips and Seeing Voices." *Nature* 264: 746–48.

McMillan, James B., and Michael B. Montgomery. 1989. *Annotated Bibliography of Southern American English*. Tuscaloosa: Univ. of Alabama Press.

Munro, Murray J., Tracey M. Derwing, and James E. Flege. 1999. "Canadians in Alabama: A Perceptual Study of Dialect Acquisition in Adults." *Journal of Phonetics* 27: 385–403.

Nábělek, Anna K., and Alexandra Ovchinnikov. 1997. "Perception of Nonlinear and Linear Formant Trajectories." *Journal of the Acoustical Society of America* 101: 488–97.

Nairn, Moray J., and James R. Hurford. 1995. "The Effect of Context on the Transcription of Vowel Quality." In *Studies in General and English Phonetics: Essays in Honour of Professor J. D. O'Connor*, ed. Jack Windsor Lewis, 96–120. London: Routledge.

Nearey, Terrance M. 1989. "Static, Dynamic, and Relational Properties in Vowel Perception." *Journal of the Acoustical Society of America* 85: 2088–113.

Niedzielski, Nancy. 1999. "The Effect of Social Information on the Perception of Sociolinguistic Variables." *Journal of Language and Social Psychology* 18: 62–85.

Ohala, John J. 1981. "The Listener as a Source of Sound Change." In *Papers from the Parasession on Language and Behavior, Chicago Linguistic Society, May 1–2, 1981*, ed. Carrie S. Masek, Roberta A. Hendrick, and Mary Frances Miller, 178–203. Chicago: Chicago Linguistic Society.

———. 1985. "Linguistics and Automatic Processing of Speech." In *New Systems and Architectures for Automatic Speech Recognition and Synthesis*, ed. Renato De Mori and Ching Y. Suen, 447–75. Berlin: Springer.

———. 1989. "Sound Change Is Drawn from a Pool of Synchronic Variation." In *Language Change: Contributions to the Study of Its Causes*, ed. Leiv Egil Breivik and Ernst Håkon Jahr, 173–98. Berlin: de Gruyter.

———. 1993. "The Phonetics of Sound Change." In *Historical Linguistics: Problems and Perspectives*, ed. Charles Jones, 237–78. Longman Linguistics Library. London: Longman.

Ohala, John J., and Deborah Feder. 1994. "Listeners' Normalization of Vowel Quality Is Influenced by 'Restored' Consonantal Context." *Phonetica* 51: 111–18.

Pear, Tom Hatherley. 1931. *Voice and Personality as Applied to Radio Broadcasting*. New York: Wiley.

Peeters, Wilhelmus Johannes Maria. 1991. "Diphthong Dynamics: A Cross-Linguistic Perceptual Analysis of Temporal Patterns in Dutch, English, and German." Ph.D. diss., Rijksuniversiteit te Utrecht.

Pitermann, Michael. 2000. "Effect of Speaking Rate and Contrastive Stress on Formant Dynamics and Vowel Perception." *Journal of the Acoustical Society of America* 107: 3425–37.

Plichta, Bartek. 2001. "Hearing Faces: The Effects of Ethnicity on Speech Perception." Paper presented at the 30th conference on New Ways of Analyzing Variation (NWAV 30), Raleigh, N.C., 11–14 Oct. Available from http://bartus.org.

Plichta, Bartek, and Norma Mendoza-Denton. 2001. "Acquisition, Processing, and Analysis of Acoustic Speech Signals." Paper presented at the 30th conference on New Ways of Analyzing Variation (NWAV 30), Raleigh, N.C., 11–14 Oct. Available from http://bartus.org.

Preston, Dennis R. 1986. "Five Visions of America." *Language in Society* 15: 221–40.

———. 1993a. "Folk Dialectology." In *American Dialect Research*, ed. Dennis R. Preston, 333–77. Amsterdam: Benjamins.

———. 1993b. "Two Heartland Perceptions of Language Variety." In *"Heartland" English: Variation and Transition in the American Midwest*, ed. Timothy C. Frazer, 23–47. Tuscaloosa: Univ. of Alabama Press.

———. 1996. "Where the Worst English Is Spoken." In *Focus on the USA*, ed. Edgar W. Schneider, 297–360. Varieties of English around the World, General Series, vol. 16. Amsterdam: Benjamins.

———, ed. 1999. *Handbook of Perceptual Dialectology*. Vol. 1. Amsterdam: Benjamins.

Purnell, Thomas, William Idsardi, and John Baugh. 1999. "Perceptual and Phonetic Experiments on American English Dialect Identification." *Journal of Language and Social Psychology* 18: 10–30.

Roberts, Margaret M. 1966. "The Pronunciation of Vowels in Negro Speech." Ph.D. diss., Ohio State Univ.

Rundell, Edward Eugene. 1973. "Studies of the Comprehension of Black English." Ph.D. diss., Univ. of Texas at Austin.

Ryan, Ellen Bouchard. 1979. "Why Do Low-Prestige Varieties Persist?" In *Language and Social Psychology*, ed. Howard Giles and Robert N. St. Clair, 145–57. Oxford: Blackwell.

Sebastian, Richard J., and Ellen Bouchard Ryan. 1985. "Speech Cues and Social Evaluation: Markers of Ethnicity, Social Class, and Age." In *Recent Advances in Language, Communication, and Social Psychology*, ed. Howard Giles and Robert N. St. Clair, 112–43. London: Erlbaum.

Shuy, Roger W. 1970. "Subjective Judgments in Sociolinguistic Analysis." In *Report of the 20th Annual Round Table Meeting on Linguistics and Language Studies*, ed. James E. Alatis, 175–88. Washington, D.C.: Georgetown Univ. Press.

Shuy, Roger W., Joan C. Baratz, and Walter A. Wolfram. 1969. *Sociolinguistic Factors in Speech Identification.* National Institute of Mental Health Research Project No. MH 15048-01. Washington, D.C.: Center for Applied Linguistics.

Smith, Bruce L., Bruce L. Brown, William J. Strong, and Alvin C. Rencher. 1975. "Effects of Speech Rate on Personality Perception." *Language and Speech* 18: 145–52.

Stephan, Christoph. 1997. "The Unknown Englishes? Testing German Students' Ability to Identify Varieties of English." In *Englishes around the World: Studies in Honour of Manfred Görlach*, vol. 1, *General Studies, British Isles, North America*, ed. Edgar W. Schneider, 93–108. Varieties of English around the World, General Series 18. Amsterdam: Benjamins.

Strand, Elizabeth A. 1999. "Uncovering the Role of Gender Stereotypes in Speech Perception." *Journal of Language and Social Psychology* 18: 86–100.

Strange, Winifred. 1989. "Evolving Theories of Speech Perception." *Journal of the Acoustical Society of America* 85: 2081–87.

Strange, Winifred, and Ocke-Schwen Bohn. 1998. "Dynamic Specification of Coarticulated German Vowels: Perceptual and Acoustical Studies." *Journal of the Acoustical Society of America* 104: 488–504.

Stroud, Robert Vernon. 1956. "A Study of the Relations between Social Distance and Speech Differences of White and Negro High School Students of Dayton, Ohio." Master's thesis, Bowling Green State Univ.

Surprenant, Aimée, and Louis Goldstein. 1998. "The Perception of Speech Gestures." *Journal of the Acoustical Society of America* 104: 518–29.

Sussman, Joan E. 2001. "Vowel Perception by Adults and Children with Normal Language and Specific Language Impairment: Based on Steady States or Transitions?" *Journal of the Acoustical Society of America* 109: 1173–80.

Thomas, Erik R. 2000. "Spectral Differences in /ai/ Offsets Conditioned by Voicing of the Following Consonant." *Journal of Phonetics* 28: 1–25.

Thomas, Erik R., and Jeffrey Reaser. 2002. "Perceptual Cues Used for Ethnic Labeling of Hyde County, North Carolina, Voices." Paper presented at the annual meeting of the American Dialect Society, San Francisco, 3–5 Jan.

Traill, Anthony, Martin J. Ball, and Nicole Müller. 1995. "Perceptual Confusion between South African and British English Vowels." In *Proceedings of the 13th International Congress of Phonetic Sciences, ICPhS 95: Stockholm, Sweden, 13–19 August 1995*, ed. Kjell Elenius and Peter Branderud, 620–23. Stockholm: Royal Institute of Technology and Stockholm Univ.

Trent, Sonja A. 1995. "Voice Quality: Listener Identification of African-American versus Caucasian Speakers." *Journal of the Acoustical Society of America* 98: 2936.

Trudgill, Peter. 1983. *On Dialect: Social and Geographical Perspectives.* New York: New York Univ. Press.

Tucker, G. Richard, and Wallace E. Lambert. 1969. "White and Negro Listeners' Reactions to Various American-English Dialects." *Social Forces* 47: 463–68.

Underwood, Gary N. 1974. "How *You* Sound to an Arkansawyer." *American Speech* 49: 208–15.

Verbrugge, Robert R., and Brad Rakerd. 1986. "Evidence of Talker-Independent Information for Vowels." *Language and Speech* 29: 39–57.

Walton, Julie H., and Robert F. Orlikoff. 1994. "Speaker Race Identification from Acoustic Cues in the Vocal Signal." *Journal of Speech and Hearing Research* 37: 738–45.

Williams, Angie, Peter Garrett, and Nikolas Coupland. 1999. "Dialect Recognition." In Preston, 345–58.

Willis, Clodius. 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: 246–55.

Wolfram, Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech.* Washington, D.C.: Center for Applied Linguistics.

———. 2000. "On Constructing Vernacular Dialect Norms." In *Chicago Linguistic Society 36: The Panels,* ed. Arika Okrent and John Boyle, 335–58. Chicago: Chicago Linguistic Society.

Wolfram, Walt, Kirk Hazen, and Natalie Schilling-Estes. 1999. *Dialect Change and Maintenance on the Outer Banks.* Publication of the American Dialect Society 81. Tuscaloosa: Univ. of Alabama Press.

Yang, Byunggon. 1992. "An Acoustical Study of Korean Monophthongs Produced by Male and Female Speakers." *Journal of the Acoustical Society of America* 91: 2280–83.