Volume 15
Issue 2 *Selected papers from NWAV 37*

Article 6

3-21-2010

What do Listeners Know about Sociolinguistic Variation?

Laura Staum Casasanto
Max Planck Institute for Psycholinguistics, laura.casasanto@mpi.nl

What do Listeners Know about Sociolinguistic Variation?

Abstract
A pair of experiments addressed two questions regarding listeners' sociolinguistic knowledge: First, how do listeners use facts about speech to inform their beliefs about speakers? Second, and conversely, how might listeners use facts about speakers to inform their perceptions of speech? Results of Experiment 1 demonstrate that listeners can infer characteristics of speakers from their use of an individual sociolinguistic variable. Results of Experiment 2 show that listeners use social information about speakers to understand ambiguous speech. Together, these results show bidirectional influences between language processing and the process of social inferencing: information in the speech stream affects inferences about social characteristics of the speaker, and social information affects speech perception.

What do Listeners Know about Sociolinguistic Variation?

Laura Staum Casasanto*

1 Introduction

This paper focuses on a pair of questions that is at the intersection of sociolinguistics and psycholinguistics: How do listeners use facts about speech to make inferences about speakers? And how might they also use facts about speakers to inform their perceptions of speech?

Linguistic variation is structured by a wide variety of social and stylistic factors. The discovery of this structure launched the field of sociolinguistics, and demonstrated that the systematic study of variable linguistic behavior is both possible and fruitful. With respect to speech perception, the importance of this discovery (and the work that has followed it) is that it uncovered a vast system of statistical regularity between language and contextual factors that could potentially be exploited by listeners for the purpose of making inferences about both language and context.

Sociolinguistics has focused traditionally on speech production, but the role of the listener is also important in explaining the systematic nature of sociolinguistic variation. Variable speech production can only be meaningful if it involves the transmission of information from speaker to listener. Information transmission requires a listener who is able to receive the information and interpret its meaning. If listeners are not at least subconsciously aware of sociolinguistic variation, sociolinguists are faced with a difficult problem: how can speakers/listeners develop such consistent and intricate patterns of sociolinguistically conditioned production if they are not influenced by each other's linguistic behavior? Thus, theories of the social meaning of linguistic variables rely crucially on the role of the listener.

The question of how listeners interpret sociolinguistic variation has been investigated with regard to a few linguistic variables and social situations. In one of the earliest experiments, Labov (1966) played samples of speech containing five socially-stratified phonological variables to listeners in New York, who evaluated the probable occupation of the speaker of each sample. The speakers' use of the socially stratified phonological variables strongly predicted the occupations that listeners assigned to them, with speakers who used higher-status variants of the phonological variables being assigned to occupations requiring standard language use. The realization of the final nasal in (ING) also influences listeners' judgments about the person who used it and the speech situation in which it occurred. For example, the alveolar nasal /n/ makes speech sound more casual and less educated and articulate, while the velar nasal /ŋ/ makes speech sound more formal and more educated and articulate (Campbell-Kibler 2007).

However, very little is known about the relationships between the production and the perception of variable linguistic behavior. To what extent do listeners store information about the structure that linguists have observed in socially conditioned variable production? If listeners monitor this information, how do they use the knowledge they accumulate?

To find out how listeners use their knowledge of sociolinguistic variation, the experiments presented in this paper use the linguistic variable known as *t/d* deletion as a testbed. This is a phonetic variable in English, also known as consonant cluster reduction, in which final coronal stops in consonant clusters may be deleted in some environments (*fas(t) car; ban(d) practice*). Final *t/d* deletion is defined as the absence of a pronounced oral stop segment corresponding to a final *t* or *d* in words (Gregory et al. 1999). The possible realizations of the final consonant vary along a continuum from an aspirated *t* with a strong release burst to a completely deleted *t,* which

*This work was supported by a Mellon Foundation dissertation fellowship to the author, and a Stanford Graduate Research Opportunity grant to the author. The work benefited from useful discussions with D. Casasanto, E.E. Coppock, K. Drager, P. Eckert, R.D. Greene, L. Hall-Lew, S. Lewis, B. Pajak, J.R. Rickford, R.L. Starr, and A. Zwicky. A preliminary report of some of the results described here was made in the proceedings of BLS 34.

leaves few spectral remnants in the acoustic signal.

Consonant cluster reduction is conditioned by several social characteristics of speakers, including race. In particular, Black speakers delete final stops more often than White speakers do (Rickford 1999, Wolfram 1969a). Because it is a well-studied feature of Black speech, *t/d* deletion has sometimes been described as a feature of African American Vernacular English (commonly referred to as *AAVE*) (Fasold 1972, Wolfram 1969a). However, an important finding of Wolfram's 1969a study of Black speakers in Detroit, among others involving *t/d* deletion, is that upper-middle class African Americans have very high rates of *t/d* deletion (nearly 80% in some phonological environments), even if they use very few other dialect features of AAVE (Wolfram 1969a). This suggests that although *t/d* deletion may be a feature of AAVE, it is also a feature of non-vernacular Black Englishes. Because this variable is associated with the race of speakers, listeners could potentially use information about *t/d* deletion to make inferences about this social characteristic. Experiment 1 investigated whether the association between *t/d* deletion and race is part of listeners' sociolinguistic knowledge. Experiment 2 investigated what kinds of inferences listeners make on the basis of this knowledge.

2 Experiment 1

Experiment 1 investigated the inferences that listeners make about social characteristics of speakers based on their use of sociolinguistic variables. Are speakers able to associate the use of a single variable with the social characteristics that condition the production of the variable? If so, do listeners base their inferences on correlations that they have observed between the social characteristics of speakers and features of the speech sounds they produce? Or, alternatively, are these associations mediated by stereotypes, which may not be grounded in correlations between social groups and specific variables?

To find out, Experiment 1 examined whether listeners attribute *t/d* deletions differently to speakers of different races, and whether different attributions are based on the distribution of these variants in speech. In a modified Matched Guise study (Lambert et al. 1960), minimal pairs of sentences were constructed to create "guises" that differ only in one dimension. In the current study, the guises were controlled and differed only in *t/d* deletion (e.g., the presence or absence of *t* or *d* word-finally). While the technique as developed by Lambert and colleagues originally involved oral guises, written guises have also been used in this paradigm, when the variables or varieties in question permitted it (e.g., Mulac, Bradac, and House 1988, Buchstaller 2004, Kramer 1974). Written guises have the advantage of allowing greater control over some orthogonal factors such as other phonological variables. For example, in this case a written guise eliminates potential variation in the realization of vowels or other non-*t/d* consonants that might influence participants' beliefs about who the speaker might be.

During the experiment, participants saw sentences that included a word ending in either a deleted or a non-deleted consonant, and judged which of two possible speakers was likely to have said the sentence. Each sentence was paired with one White and one Black potential speaker. If listeners store information about the distribution of deleted final consonants in the input, they should rate more sentences with deleted tokens as having been spoken by a Black speaker than sentences with non-deleted tokens.

2.1 Participants

Stanford University undergraduates (N=111) received course credit in an introductory psychology class for their participation in this study, which was distributed as one page in a larger packet of unrelated surveys. Participants were of both genders and a mixture of ethnicities, and most participants were between the ages of 18 and 22. Because the class was not in the area of linguistics, most participants presumably had little or no linguistics training and were probably not aware of the status of *t/d* deletion as a sociolinguistic variable.

2.2 Materials

Participants saw 24 experimental items and 12 control items, each comprising one sentence below two faces of people who might have spoken the sentence.

Twenty-four sentences were constructed so that each included a word with a consonant cluster that could be subject to *t/d* deletion (e.g., *mast, least, wind*). The *t/d* consonants were primarily in phonological environments that promote consonant cluster reduction, such as following a homogeneously voiced fricative or a nasal and preceding a word beginning with a stop or a glide (e.g., *The mast probably lasted…*). Each sentence appeared in two versions, the standard version and the deleted (non-standard) version. In the standard version, this word was presented with its normal orthography (e.g., *mast*). In the deleted version, this word appeared with its final stop replaced by an apostrophe (e.g., *mas'*), indicating a deleted final consonant.

Underneath each sentence in Questionnaire B, participants saw a "translation" of the non-standardism to ensure that all participants interpreted the stimuli as realizations of the same words. In the case of *t/d* deleted items, this consisted of the same word written in standard orthography (i.e., the word *mast* was written underneath the form *mas'*). Items were counterbalanced across versions of each questionnaire so that although each participant saw only twelve target items, all twenty-four target items were seen in both standard and deleted versions by some participants.

Participants received one of two questionnaires, A (standard) or B (non-standard). Each questionnaire contained twelve target items. Questionnaire A contained the standard versions of the target *t/d* deletion sentences (presented in normal orthography), and twelve similar sentences presented in normal orthography (which served as fillers). Questionnaire B contained the deleted (non-standard) versions of the target *t/d* deletion sentences (written with an apostrophe in place of the final consonant). Questionnaire B also contained twelve control items, which contained an unrelated non-standard usage. Thus, all the sentences in Questionnaire B contained a departure from Standard Written English, expressed orthographically, whereas all sentences in Questionnaire A contained sentences in Standard Written English.

Four types of non-standardisms were used, with three items of each type to produce the twelve items. These other types of non-standardisms also appeared more than once so that *t/d* deletion was not unique in appearing multiple times in the questionnaire.

One of the four types of other non-standardisms was another phonetic variant, the extreme diphthong in the New York pronunciation of words like *coffee* and *dog* ([kɔ^{ə}fi], [dɔ^{ə}g]); these vowels were represented orthographically as *cawfee* and *dawg*. Like the *t/d* deletion items, these were "translated" below the sentences by the standard orthography of the words, to ensure that the fillers were as similar as possible to the target items. Two of the other non-standardisms were morphosyntactic. The first was the *needs washed* construction, as in *The car parked near my house needs washed*. These were "translated" below the sentences as *needs to be washed*. The second morphosyntactic variable was double modals, as in *My twin brother might could pass for me at school.* This was "translated" as *might be able to*. The fourth non-standardism was lexical; the phrase *youse guys* appeared in the sentences, translated as *you guys*. Importantly, none of these variables is used more by Black speakers than by White speakers in the same speech community (Labov 1966, Murray, Frazer, and Simon 1996, Smitherman 1986).

The term "non-standard" is used here as an umbrella term, encompassing different types of usages. The four "other non-standard phenomena" differ from each other in the way they depart from an imagined national standard: some are better described as vernacular, some are better described as regionalisms, and others as both. The important way in which these usages form a natural class with *t/d* deletion is that they are noticeably different from a "standard" alternative. While it is hard to know what constitutes the imagined national standard of our participants, all the usages included in Questionnaire B carry less global prestige than most of their alternatives.

The other non-standardisms were selected because they are not associated with Black speakers in production. However, the control items differed from *t/d* deletion in other ways, as well. Most notably, they included some lexical and morphosyntactic non-standard features. Because not

all non-standard features are equally marked, it could be that these features were less noticeable to participants than *t/d* deletion (a phonetic non-standardism), and the sentences containing these lexical and morphosyntactic features might seem less non-standard overall. If so, *t/d* deletion could be associated more strongly with Black speakers due to exactly the stereotyping effect that the control condition is intended to rule out. However, language attitude surveys have shown that non-standard lexical and/or morphosyntactic features tend to be more sharply stratified among speakers, and are more noticeable to listeners than non-standard phonetic features are (Lippi-Green 1997, Wolfram 1969b). Thus, stereotype-based judgments driven by the relative markedness or noticeability of the target and control items would likely result in the control sentences being more strongly associated with Black speakers, contrary to the experience-based prediction.

In summary, each questionnaire had 24 sentences; Questionnaire A contained 12 standard (non-deleted) experimental items and 12 standard fillers, and Questionnaire B contained 12 non-standard (deleted) experimental items and 12 non-standard fillers.

Above each sentence on both questionnaires were two pictures. Pictures of potential speakers were taken from a database of University of Pennsylvania ID photos (Killgore et al. 2000), and included the shoulders and head of college-aged individuals, on a white or neutral background. Four Black and four White male individuals were used for both target and filler items, with a variety of pairings of Black and White individuals. The photographs were restricted to males because males tend to have higher rates of *t/d* deletion overall than females do. Using photographs of males for the potential speakers made the deleted versions of the sentences more plausible, and also eliminated the problem of cross-gender pairings in which gender might play as much of a role (if not more) as race in determining participant responses. Each sentence had one White and one Black individual pictured above it.

2.3 Procedure

Each participant was randomly assigned to receive one of the two questionnaires. In both questionnaires, their task was to circle the picture of the person they thought was more likely to have said the sentence in question. The instructions for each questionnaire varied, as below:

> Questionnaire A instructions:
> *Below each pair of pictures is a sentence. For each pair, please circle the person you think is more likely to have said the sentence.*
>
> Questionnaire B instructions:
> *The sentences below are transcribed from natural speech. Some of the words speakers used are non-standard. "Translations" for these words are given in parentheses. Try to imagine how they might have sounded in your mind's ear. For each sentence, decide which person pictured above it is more likely to have said it.*

Fifty-eight participants received Questionnaire A; the remaining 53 participants received Questionnaire B. Using written stimuli allowed this experiment to address the influence of *t/d* deletion without the influence of auditory cues to race. Because participants' association between *t/d* deletion and race might not be conscious, participants were not made aware that either consonant cluster reduction or race was of interest. Other non-standardisms were present in the fillers to mask the target variable. However, due to the fact that all pairs of faces contained one White and one Black face, participants probably realized that race was of interest in the experiment.

Because the standard versions of the sentences appeared only in Questionnaire A and the deleted versions appeared only in Questionnaire B, the difference between standard and deleted versions was a between-subjects comparison. This prevented sentences containing *t/d* deletions from standing out as especially non-standard to any given participant, and prevented participants from being able to compare deleted vs. non-deleted versions of similar sentences.

2.4 Results and Discussion

Participants attributed 60% of the "deleted" target sentences, represented with apostrophes, to the pictures of Black students pictured, and the remaining 40% to the pictures of White students. By contrast, they attributed only 42% of the non-deleted target sentences, with normal orthography, to the pictures of Black students, attributing the remaining 58% to those of the White students. This pairwise difference between proportions was significant (t1(1,109)=4.86, $p<.001$; t2(1,46)=3.98, $p=.0002$; minF'(1, 108=9.49, $p=.003$), indicating that participants do associate *t/d* deletion more with Black speakers than with White speakers.

But how specific is listeners' knowledge of links between language and race? The finding that speakers associate race with *t/d* deletion does not entail that listeners have attended to variation in the world and stored information about it. Rather, it is possible that this association could be due to the application of more general stereotypes. The term *stereotype* refers to beliefs that can be based on things like hearsay and media representations, rather than direct experience with people. Because it's possible to hold beliefs about how people of different ethnicities act without actually having observed them acting in this manner, evidence that a belief exists is not necessarily informative about how the belief was acquired.

Could stereotyping, rather than direct experience with speakers of different ethnicities, be responsible for the results of Experiment 1? In principle, this is plausible. Because the deleted variant is less standard than the non-deleted variant, participants could have attributed *t/d* deletion to Black speakers simply because they associate less standard usages with Black speakers. This would not require any experience on the part of the listener with the specific variable, only the belief that 1) reductions and deletions are non-standard, and 2) Black speakers are more likely to produce non-standard language.

If participants associate less standard usage with Black speakers based only on stereotyping, then they should show the same preference for selecting Black speakers with other non-standardisms that they showed for *t/d* deletion, even if the other non-standardism is not associated with Blacks in production. However, if they associate *t/d* deletion more with Black speakers than White speakers based on experience with the variable itself, then their bias should be selective for *t/d* deletion, because Black speakers produce *t/d* deletion more than their White counterparts.

Results showed that control sentences containing the other non-standardisms were attributed to Black speakers 51% of the time. This result was no different from chance, which is 50%.[1] Importantly, the results of the control sentences were significantly different from the results of the target sentences, in which 60% of *t/d* deleted sentences were attributed to Black speakers (t1(1,52)=1.97, $p=.03$; t2(1,46)=2.27, $p=.02$; minF'(1,97)=2.22, $p=.14$). This pattern suggests that participants associated Black speakers with *t/d* deletion in particular, and not just with non-standard usages in general.

2.5 Summary of Experiment 1

Experiment 1 demonstrated that listeners associate *t/d* deletion more strongly with Black speakers than with White speakers, which is consistent with the facts about production of this variable. Results of the control condition indicate that although stereotyping may play a role in forming this association, listeners reliably distinguish between *t/d* deletion and other non-standard variants that they have likely never heard a Black speaker produce. This suggests an important role for direct experience with linguistic material (via face-to-face contact or recorded media) in forming the associations listeners have between social characteristics and linguistic variants.

[1]One item that was intended to represent a New York City vowel (*dawg*) was associated strongly with Black speakers by participants, presumably because this spelling has other interpretations besides the intended one. Removing this item from the analysis very slightly decreases the percentage of other non-standardisms attributed to Black speakers (to 50%) and does not change the relationship between conditions.

3 Experiment 2

Experiment 1 confirmed that listeners have access to statistical relationships between *t/d* deletion and race, and suggested that language experience is involved in accumulating knowledge of these relationships. Yet, this experiment leaves an important question unanswered: do listeners use this knowledge of sociolinguistic variation to help them understand language?

Previous studies have established that listeners can use social information presented visually or in explicit instructions to participants, as well as social information that they must gather from an auditory signal, to inform their expectations about how speakers will produce different phones (Glidden and Assmann 2004, Hay, Nolan, and Drager 2006, Johnson, Strand, and D'Imperio 1999, Niedzielski 1999, Plichta and Rakerd 2002). These effects have been observed in phoneme categorization tasks, which involve explicit decisions about the linguistic variable in question. However, listeners should also make use of social information to resolve naturally occurring ambiguity in on-line language comprehension.

3.1 Participants

Native U.S. English speakers (N=44) from the Stanford University community participated in this study in exchange for payment. Participants were of a variety of ethnicities and both genders.

3.2 Materials

Twenty-four pairs of sentences were constructed which were identical for the first few words (the section underlined in 1a and 1b below) except for a critical word (italicized below). The critical words in each pair of sentences were identical except for the presence or absence of a stop at the end of a final consonant cluster:

(1) a. **The *mast* probably lasted** through the storm.
b. **The *mass* probably lasted** an hour on Sunday.

These nearly identical sections (in bold above) would be ambiguous when spoken aloud if a speaker used the deleted variant of a word like *mast*. Because the deleted variant is a possibility, a listener would not be able to tell whether the word *mass* or *mast* was intended at this point in the sentence. The pairs of sentences, however, are all disambiguated by the endings of the sentences, which are more consistent with one of the interpretations of the beginning. For example, *through the storm* is more consistent with the *mast* interpretation of the beginning, and *an hour on Sunday* is more consistent with the *mass* interpretation of the beginning. The sound files used in this experiment were only excerpted from recordings of the sentence in each pair that never contained a final stop (e.g., the *mass* version). Participants never heard any version of the experimental sentences in which the target word was intended to contain an underlying *t/d* by the speaker, so that there were no cues in the recordings indicating the presence of a deleted stop.

Twenty-four filler pairs were created that also contained an ambiguity that was resolved later in the sentences:

(2) a. While Bill hunted the deer ran into the woods.
b. While Bill hunted the deer we made the fire.

None of these ambiguities were related to *t/d* deletion. While some of these temporary ambiguities could be disambiguated by prosody, care was taken to select recordings in which the prosody would be appropriate for both readings of the ambiguous portion. These ambiguous fillers masked the experimental sentences; when a subject encountered a temporarily ambiguous sentence, it was only an experimental sentence half the time. This prevented the experimental sentences from

standing out from the fillers, as well as preventing them from exclusively constituting the most difficult sentences to understand and respond to.

In addition, 48 unambiguous filler sentences were constructed of similar length and complexity. Thus the full set of stimuli includes 24 pairs of experimental stimuli, 24 pairs of ambiguous fillers, and 48 unambiguous fillers, which were unpaired. Because each subject heard only the ambiguous portion of each pair, this resulted in 96 total sentence beginnings presented to each subject, half of which were ambiguous. The ratio of fillers to target stimuli was 3 to 1.

Sixteen Stanford graduate and undergraduate students were recorded while reading these sentences aloud. The participants were 4 Black males, 4 White males, and 8 females of various races/ethnicities. Each sentence was preceded by a context sentence to make the participants' reading of the experimental and filler sentences more natural:

(3) a. I went to a new church last week that has very short services. (context)
b. The mass probably lasted an hour on Sunday. (target)

Many of these speakers were recruited from the Stanford University Linguistics Department subject pool; because the subject pool did not contain many Black participants, additional Black participants were recruited via email and paid ten dollars for their time. Recordings were digital, made in a sound-attenuated booth and displayed on a computer screen.

Recordings of the male speakers were used for the target stimuli and some of the fillers. Recordings of the female speakers were used for the remaining fillers (making up half of the total clips heard by participants). Each target item was recorded once by a Black male speaker, and once by a White male speaker. Participants were randomly assigned to hear an equal number of items recorded by Black and White speakers in each face condition, so that the pairing of voice and face was equally felicitous across conditions, on average.

The speaker pictures from Experiment 1 (4 Black males and 4 White males) were used for the critical trials and one third of the fillers (24 targets and 24 fillers), while 8 females of various races/ethnicities were displayed with the other two thirds of the fillers (48). This resulted in each subject seeing a female face in half the trials and a male face in the other half of the trials. The faces and voices were paired so that each face appeared with only one voice within each subject, to maximize the plausibility of the premise that the pictures represented the speakers. Between subjects, each face appeared with one Black voice and one White voice, so the influence of the faces can be evaluated independently of the influence of voices. Male voices always appeared with male pictures, and female voices always appeared with female pictures.

Because all subjects saw a mix of genders and races/ethnicities in the experiment, they were unlikely to be able to deduce that the experimental conditions concerned only Black and White males. Thus, unlike in Experiment 1, neither *t/d* deletion nor race was salient in this experiment.

3.3 Procedure

Participants were instructed to listen to a short sound clip while looking at a picture of a face, which they were told represented the speaker of the clip. They heard the ambiguous portion of one of the sentence pairs, which contained no final stop at the end of the cluster in the target word. After hearing the beginning of the sentence, participants then saw one of the sentence endings appear below the picture of the speaker. For example, in one trial, participants heard:

(4) The [mæs] probably lasted

After this clip, one of the following endings appeared on the screen:

(5) a. …through the storm.
b. …an hour on Sunday

In half the cases, participants saw a continuation that was more likely if the ambiguous word had no final stop (e.g., *an hour on Sunday*, which was more likely if the word was *mass*), and in the other half of cases they saw the other continuation, which was more likely if the ambiguous word did have a final stop that had been deleted (e.g., *through the storm*, which was more likely if the word was *mast*). Participants' job was to assess whether the ending created a "sensible" sentence in combination with the beginning they had heard, and response times were measured from the time the continuation appeared on the screen.

Because the ambiguous words could be interpreted as either *mass* or *mast* and thus both continuations could be interpreted as sensible, very few *no* responses were expected. Since a *yes/no* response measure was unlikely to be sensitive enough to detect an effect of race, response time was the dependent measure in this design. Reaction times were measured from the point when the face appeared on the screen and the audio file of the sentence beginning started until the point that the participant pressed the Y or N key to respond. If listeners use the relationship between race and *t/d* deletion in resolving ambiguity, then they should respond faster to the continuation that is consistent with the *t*-word (*mast*) interpretation when the purported speaker is Black than when he is White. Conversely, listeners should respond faster to the continuation that is consistent with the non-*t*-word (*mass*) interpretation when the purported speaker is White than when he is Black.

3.4 Results and Discussion

Two participants who were non-native speakers of English were excluded. Two more participants were excluded because their average reaction times were more than 2 standard deviations higher than the mean, leaving 40 participants in the analysis. Reaction times longer than 10 seconds were removed from the analysis (<1% of trials).

A 2-way ANOVA with Race (Black, White) and Deletion (deleted, non-deleted) as between-subject factors showed a significant interaction in the predicted direction (deleted *t/d*, non-deleted *t/d*; ($F1(1,38)=5.64$, $p=.02$, $F2(1,22)=9.23$, $p=.006$; min$F'(1,60)=3.5$, $p=.066$). Post-hoc comparisons indicated that participants responded faster to the continuation that was compatible with the word whose underlying phonemic form has a *t* (the *mast* interpretation) when they saw a Black face (t1(1,39)=1.21, p=.11, t2(1,23)=1.8, $p=.04$; min$F'(1.59)=1.01$, $p=.32$). However, they showed the opposite pattern (responding faster to the continuation that was compatible with the word whose underlying phonemic form does not have a *t,* the *mass* interpretation) when they saw a White face (t1(1,39)=1.75, $p=.04$, t2(1,23)=1.81, $p=.04$; min$F'(1,57)=1.58$, $p=.21$).

The results of Experiment 2 suggest that listeners make use of sociolinguistic variation in language processing. However, interpreting this pattern of results as supporting a role for social information in speech comprehension requires that the presence or absence of a final stop be the only important difference between the *t*-word sentence endings and the non-*t*-word sentence endings. The two sentence endings in each pair differ in several ways beyond the fact that some of them implied an instance of *t/d* deletion in the beginning of the sentence and some did not: the content of the sentence endings is different, and the register of the sentences could be different, among other things. Thus, there are many possible differences between the sentence endings that could result in one continuation being responded to more quickly than the other in each pair, beyond participants' linguistic stereotypes about *t/d* deletion. Differences in the continuations that are unrelated to *t/d* deletion could have caused the pattern of results observed in Experiment 2 if they make the *t*-word endings more associated with Black speakers, on average, than the non-*t*-word endings. This association could be based on meaning differences in the sentences; for example, in the *mass/mast* pair of sentences, one sentence is about church-going, and the other is about ships. If participants associated church-going more with Black speakers and ships more with White speakers, then the association between these activities and race would predict a difference in reaction time to the two sentence endings. Because results in Experiment 2 were aggregated across items, they could only be explained by this kind of association if the *t*-word sentence endings were on average more associated with Black speakers than the non-*t*-word sentence endings were. If so,

the results of Experiment 2 could be attributable to the content of the sentences, rather than the need to restore a deleted *t* in the *t*-word cases.

The *t*-word sentences could also be associated with Black speakers due to register or dialect differences from the non-*t*-word sentences (although any differences must reside in the written sentence endings, so phonological factors cannot be involved). For example, if the *t*-word sentence endings contained lexical items or syntactic constructions that are characteristic of some variety of African American English, this could cause them to be associated with speakers of that variety.
To rule out these alternative interpretations, I conducted an additional analysis on the data from Experiment 1, in which participants assigned the written versions of the same sentences to a likely speaker. These written versions have all the same content, register and dialect differences that the sentences used in Experiment 2 have, but they are unambiguous with respect to *t/d* deletion. In Experiment 1, participants were asked to select which of the two pictures above each sentence was more likely to represent the speaker of the sentence. Each participant saw only one sentence from each pair, so that no two sentences started with the same words for each participant. Thus, the comparison between the two sentences in each pair is a between-subjects comparison.

There was no significant difference between the *t*-word and non-*t*-word sentences ($t1(1,57)=1.07$, $p=.29$; $t2(1,46)=.84$, $p=.40$; $\min F'(1, 90)=.43$, $p=.51$). In addition, the slight numerical difference between the two lists suggests that if anything, the non-*t*-word sentences were more associated with the Black potential speakers (46% of responses as opposed to 43% of responses to the *t*-word sentences). If this association were to achieve significance with increased power, it would be in the opposite direction from the difference that would be needed to account for the results of Experiment 2. Thus, general differences in content, register, or dialect between the two lists of sentences cannot account for the results of Experiment 2.

3.5 Summary of Experiment 2

Results of Experiment 2 were consistent with the experimental hypothesis that listeners use social information to resolve ambiguities, but this experiment left open the question of whether other differences in the stimuli could have caused the observed difference. The analysis of data from Experiment 1 ruled out content-, register- or dialect-based explanations for the results of Experiment 2, confirming that expectations about *t/d* deletion are the likely explanation for the observed differences in behavior. Experiment 2 provides evidence that listeners combine their knowledge of sociolinguistic variation, social information from the scene, and auditory information from the speech stream to construct an interpretation of the speech they have heard.

4 Conclusion

Results of two experiments indicate that listeners learn probabilistic relationships between social characteristics of speakers and their linguistic behavior, and use this information during language processing. Experiment 1 showed that listeners keep track of sociolinguistic variation, and use this information to make inferences about speakers. Conversely, Experiment 2 showed that listeners use social information about speakers to make inferences about speech. Making use of social cues to phonetic variation is an efficient use of available information to solve the difficult problem of creating meaning from speech.

References

Buchstaller, Isabelle. 2004. Putting perception to the reality test: The case of “go” and “like.” *U. Penn Working Papers in Linguistics 10.2: Selected Papers from NWAVE 32*, ed. M. Baranowski and U. Horesh, 61–76.

Campbell-Kibler, Kathryn. 2007. Accent, (ING), and the social logic of listener perceptions. *American*

Speech 82:32–64.
Fasold, Ralph. 1972. *Tense Marking in Black English: A Linguistic and Social Analysis.* Arlington, VA: Center for Applied Linguistics.
Glidden, Catherine, and Peter F. Assmann. 2004. Effects of visual gender and frequency shifts on vowel category judgments. *Acoustics Research Letters Online* 5:132–138.
Gregory, Michelle, William D. Raymond, Alan Bell, Eric Fossler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. In *Proceedings of CLS 35*, 151–166.
Hay, Jennifer, Aaron Nolan, and Katie Drager. 2006. From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review* 23:351–379.
Johnson, Keith, Elizabeth Strand, and Mariapaola D'Imperio. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of Phonetics* 27:359–384.
Killgore, William, Daniel Casasanto, Deborah Yurgelun-Todd, Joseph Maldjian, and John Detre. 2000. Associative memory for faces activates the left amygdala and hippocampus. *NeuroReport* 11:2259–2263.
Kramer, Cheris. 1974. Folk linguistics: Wishy-washy mommy talk. *Psychology Today* 8:82–85.
Labov, William. 1966. *The Social Stratification of English in New York City.* Washington, DC: Center for Applied Linguistics.
Lambert, Wallace E., Richard C. Hodgson, Robert C. Gardner, and Samuel Fillenbaum. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60:44–51.
Lippi-Green, Rosina. 1997. *English with an Accent: Language, Ideology and Discrimination in the United States.* New York: Routledge.
Mulac, Anthony, James Bradac, and Ann House. 1988. Lexical diversity and magnitude of convergent versus divergent style shifting: Perceptual and evaluative consequences. *Language and Communication* 8:213–228.
Murray, Thomas E., Timothy C. Frazer, and Beth L. Simon. 1996. Need + past participle in American English. *American Speech* 71:255–271.
Niedzielski, Nancy A. 1999. The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology* 18:62–85.
Plichta, Bartek, and Brad Rakerd. 2002. Perceptions of /a/-fronting across two Michigan dialects. Paper presented at NWAV 31.
Rickford, John R. 1999. *African American Vernacular English: Features, Evolution, Educational Implications.* Oxford: Blackwell.
Smitherman, Geneva. 1986. *Talkin and Testifyin: The Language of Black America.* Detroit: Wayne State University Press.
Wolfram, Walt. 1969a. *A Sociolinguistic Description of Detroit Negro Speech.* Washington, DC: Center for Applied Linguistics.
Wolfram, Walt. 1969b. Sociolinguistic implication for educational sequencing. In *Strategies for Teaching Standard English to Black Children*, ed. R. Fasold and R. Shuy. Washington, DC: Center for Applied Linguistics.

Max Planck Institute for Psycholinguistics
Wundtlaan 1
6525 XD, Nijmegen, NL
laura.casasanto@mpi.nl